Lecture Notes

Stochastic Integration: Itô Stochastic Integral

Maurizio Tiso

Technical Report # 633
May 27, 1999

STOCHASTIC INTEGRATION: ITÔ STOCHASTIC INTEGRAL

Maurizio Tiso

1. Introduction

Stochastic Calculus has become an indispensable tool in several fields. For istance, it is used in Mathematical Finance to price and hedge financial derivatives, in Statistics to deal with continuous-time models and survival analysis[1], in Engineering for filtering and control theory, in Physics to study the effects of random excitations on various phyisical phenomena and in Biology to model the effects of statistical variability in reproduction and environment on populations. From an applied prospective Stochastic Calculus can be loosely described as a calculus for stochastic processes or a sort of infinitesimal calculus for non differentiable functions which plays a role when the necessity of including unpredictable factors into modeling arises.

To get an appreciation of the kind of problems we are going to deal with and how they arise, it is useful to look at a couple of examples. So, if we think of M_s as the price of a stock at time s and X_s as the number of shares an investor holds, the integral $I_t = \int_0^t X_s\,dM_s$ represents the net profits at time t, relative to the wealth at time 0. To check this one should note that the infinitesimal rate of change of the integral, $dI_t = X_t\,dM_t$, equals the rate of change of the stock

[1] In the case of survival analysis it is actually the modern theory of stochastic integration and the general theory of stochastic processes rather than Itô calculus to be useful. A good reference is Fleming and Harrington (1991).

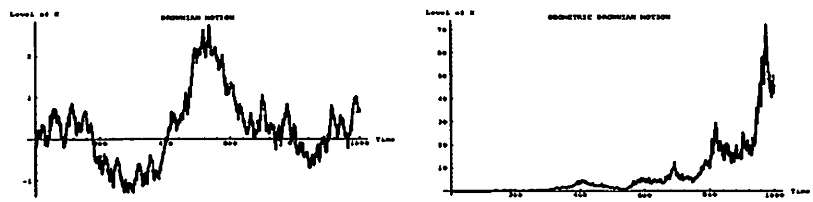

Figure 1.1: *Brownian motion and Geometric Brownian motion having the same drift ($\alpha =$ .05) and volatility ($\sigma = .3$). Both processes are assumed to start from $X_0 = .1$.*

price times the number of shares held by the investor. Clearly, the definition of I_t above is meaningful if we have available a realistic model for stock prices.
A typical choice is to assume that $M = \{M_t : t \in \mathbf{R}_+\}$ follows a geometric Brownian motion (GBM) with volatility σ and drift α, i.e.:

$$M_t = M_0 \cdot \exp\{\alpha t + \sigma W_t\}$$

which, as we will prove in a later section, translates in the following stochastic differential

$$dM_t = \left(\alpha + \frac{\sigma^2}{2}\right) M_t\, dt + \sigma M_t\, dW_t$$

where $\alpha \in \mathbf{R}$, σ is a positive real number and W is Brownian motion. The process is appropriate for economic variables that grow exponentially at an average α and have volatility proportional to the level of the variable. The process also exhibits increasing forecast uncertainty. Geometric Brownian motion is often used to model security values since the proportional changes in security price are independent and identically normally distributed. It can also be used to model anything that is parametric and increase on average at a certain exponential rate such as the nominal price of a commodity or the revenue from a particular activity. Harrison and Pliska (1981, 1983) survey the application of martingale theory and stochastic integrals to continuous trading.
Assuming a GBM as a model for the stock price we find that the profits at time t can be written as

$$I_t = \int_0^t X_s\, dM_s = \left(\alpha + \frac{\sigma^2}{2}\right) \cdot \int_0^t X_s M_s\, ds + \sigma \int_0^t X_s\, dW_s$$

and this raises several questions:

(a) What is the meaning of $\int_0^t X_s M_s\, ds$?

(b) What is the meaning of $\int_0^t X_s \, dW_s$?

(c) How do we compute these "integrals"?

The first type of integral is known as a **random integral** and it is different from standard integrals only because the integrand is a random function. The second integrals is what is referred to as a **stochastic integral**. In this case both the integrand and the measure with respect to which integration is performed are random functions. Both random and stochastic integrals are stochastic processes and, hence, are quite different from ordinary integrals, nonetheless, one does not need a special theory to compute random integrals. Stochastic integrals, on the other hand, are a different problem and will be the subject of our analysis.
As a second example one can consider the simple population growth model given by

$$\frac{dN(t)}{dt} = a(t)N(t), \quad N(0) = N_0$$

where $N(t)$ is the size of the population at time t and $a(t)$ s the relative rate of growth, it might happen that $a(t)$ is not completely known, but subject to random environmental effects, i.e.:

$$a(t) = r(t) + \sigma(t) \cdot \text{noise}.$$

Then, the growth model can be rewritten as

$$\frac{dN(t)}{dt} = r(t) \cdot N(t) + \sigma(t) \cdot N(t) \cdot \text{noise}$$

or, in form of integration,

$$N(t) = N_0 + \int_0^t r(s)N(s)ds + \int_0^t \sigma(s)N(s) \cdot \text{noise } ds.$$

This again raises questions:

(d) What is the mathematical interpretation for the noise term?

(e) What is the meaning of $\int_0^t \sigma(s)N(s) \cdot \text{noise } ds$?

It turns out that a reasonable mathematical interpretation for the noise term is the so called white noise, $WN(t)$, which is formally regarded as the derivative of a Brownian motion $W(t)$, i.e. $WN(t) = dW(t)/dt$. Thus, the noise $\cdot \, dt$ term can be expressed as $WN(t)dt = dW(t)$, and

$$\int_0^t \sigma(s)N(s) \, \text{noise} \, ds = \int_0^t \sigma(s)N(s)dW(s)$$

and, again, we find that we have to deal with stochastic integration.

The goal is then to define stochastic integrals $\int_0^t X_s\, dM_s$ also denoted $\int_{[0,t]} X\, dM$ or $(X \cdot M)_t$. If we let $\{M_n, \mathcal{F}_n : n \in \mathbf{Z}_+\}$ be a martingale and $\{X_n : n \in \mathbf{Z}_+\}$ is any process, one can define the following stochastic integral in discrete time

$$(X \cdot M)_n \equiv \sum_{m=1}^{n} X_m \cdot (M_m - M_{m-1}).$$

To this purpose if $\xi_1, \xi_2, \dots$ are independent r.v.'s with $P[\xi_i = 1] = P[\xi_i = -1] = 1/2$, letting $M_n = \xi_1 + \xi_2 + \ldots + \xi_n$, it is easily checked that M_n is a simple random walk and a martingale with respect to $\mathcal{F}_n = \sigma(\xi_1, \xi_2, \dots, \xi_n)$. If we think of a person flipping a fair coin and betting \$1 on heads each time ($X_m < 0$ meaning that the person bets $-X_m$ on tails), $(X \cdot M)_n$ gives the person's net winnings at time n. In fact, if $X_m > 0$ the gambler wins his/her bet at time n and increases his/her fortune by X_m iff $M_m - M_{m-1} = 1$.
The gambling interpretation of the stochastic integral suggests that it is natural to let the amount bet at time n depend on the outcomes of the first $n-1$ flips but not on the flip we are betting on, or on later flips.
A process $\{X_n : n \in \mathbf{Z}_+\}$ such that $X_n \in \mathcal{F}_{n-1}$ for all $n \geq 1$, ($\mathcal{F}_0 = \{\Omega, \emptyset\}$) is said to be **predictable** since its value at time n can be predicted (with certainty at time $n-1$.). The next example shows that one cannot make money by gambling on a fair game.

predictable process

Example. Let $\{M_n, \mathcal{F}_n : n \in \mathbf{Z}_+\}$ be a martingale. If $\{X_n : n \in \mathbf{Z}_+\}$ is predictable and each X_n is bounded, then $(X \cdot M)_n$ is a martingale.
Clearly, $(X \cdot M)_n \in \mathcal{F}_n$. In addition, the boundedness of the X_n's implies that $E[|(X \cdot M)_n|] < \infty$ for each n. This being established one can compute conditional expectations to conclude

$$\begin{aligned}
E[(X \cdot M)_{n+1} \mid \mathcal{F}_n] &= (X \cdot M)_n + E[X_{n+1} \cdot (M_{n+1} - M_n) \mid \mathcal{F}_n] \\
&= (X \cdot M)_n + X_{n+1} \cdot E[M_{n+1} - M_n \mid \mathcal{F}_n] \\
&= (X \cdot M)_n + X_{n+1} \cdot 0 = (X \cdot M)_n
\end{aligned}$$

since $E[M_{n+1} - M_n \mid \mathcal{F}_n] = 0$ and $X_{n+1} \in \mathcal{F}_n$ for all $n \in \mathbf{Z}_+$.

It is easy to check that the impossibility to make money by gambling on a fair game does not hold if X_n is only assumed to be **optional** (i.e. $X_n \in \mathcal{F}_n$) since we can base our bet on the outcome of the coin we are betting on.

optional process

Example. If M_n is the symmetric random walk considered above and $X_n = \xi_n$ so that $X_n \in \mathcal{F}_n$ but $X_n \notin \mathcal{F}_{n-1}$ then

$$(X \cdot M)_n = \sum_{m=1}^{n} \xi_m \cdot \xi_m = n$$

since $\xi_m^2 = 1$ for all $m \in \mathbf{Z}_+$.

Even in continuous time we would like to retain the impossibility to make money gambling on a fair game. In other words, we **want our integrals to be martingales**. This requirement can be achieved only if X is a predictable process and the next example shows that we cannot dispense with this requirement.

Example. Let $(\Omega, \mathcal{F}, P)$ be a probability space on which there is defined a r.v., T, with $P[T \leq t] = t$ for $0 \leq t \leq 1$ and and independent r.v. ξ with $P[\xi = 1] = P[\xi = -1] = 1/2$. Let

$$X_t = \begin{cases} 0 & \text{if } t < T \\ \xi & \text{if } t \geq T \end{cases}$$

and let $\mathcal{F}_t = \sigma(X_s : s \leq t)$. In other words, one waits until time T and then flips a coin. $\{X_t, \mathcal{F}_t : t \in [0,1]\}$ is a martingale. However, if we define the stochastic integral $I_t = \int_{[0,t]} X\, dX$ to be the ordinary Lebesgue-Stielties integral, then

$$Y_1 = \int_0^1 X_s\, dX_s = X_T \cdot \xi = \xi^2 = 1$$

as the measure dX_s corresponds to a mass of size ξ at T and the integral is ξ times the value there. Noting that $Y_0 = 0$ while $Y_1 = 1$, it is clearly understood that I_t cannot be a martingale.

So, we have presented an argument to support the choice of a predictable process as integrands. It is possible to define stochastic integrals when the integrators are continuous local martingales, bounded martingales, local martingales, and semimartingales. In these notes , however, we consider integration with respect to Brownian motion and its properties only. These integrals are called Itô integrals and the corresponding calculus, Itô calculus.

2. Random Integrals

Random integrals are different from usual (deterministic) integrals only because the integrand functions are actually random functions (stochastic processes). If $X = \{X_t : t \in \Gamma \subset \mathbf{R}_+\}$ is a stochastic process on some probability space $(\Omega, \mathcal{F}, P)$, a random integral is a functional like

$$I_t(\omega) = \int_0^t X_s(\omega)\, ds, \quad t \in \Gamma.$$

Example. Assume that data are generated by a random walk, i.e.:

$$y_t = y_{t-1} + \mu_t, \quad \mu_t \sim \text{i.i.d}(0, \sigma^2), \quad t = 1, 2, \ldots, T$$

but one estimates the following AR(1) model

$$y_t = \alpha + \rho y_{t-1} + \mu_t$$

using OLS. If $\hat{\rho}_T$ is the OLS estimate for ρ it is possible to prove (see e.g. Hamilton (1994), formula 17.4.28, p. 492) that when $T \to \infty$,

$$T(\hat{\rho}_T - 1) \xrightarrow{D} \frac{\frac{1}{2}(W_1^2 - 1) - W_1 \int_0^1 W_s\, ds}{\int_0^1 W_s^2\, ds - \left(\int_0^1 W_s\, ds\right)^2}$$

where W denotes Brownian motion. The asymptotic distribution for $\hat{\rho}_T$ involves two different random integrals: $\int_0^1 W_s^2\, ds$ and $\int_0^1 W_s\, ds$.

Under some conditions on the stochastic process X, random integrals are random variables on the probability space $(\Omega, \mathcal{F}, P)$. In fact, if for example we assume that $X_t(\omega)$ is continuous (right- or left-continuity is actually enough) and such that $\int_\Gamma E[X_t^2]\, dt < \infty$ and Γ is a finite interval of $\mathbf{R}_+$, this suffices to establish our assertion and, in addition, provides, enough regularity to compute the mean of the random variable $I_\Gamma(\omega) \equiv \int_\Gamma X_t(\omega)\, dt$. To see why it is so, one should notice that

- $X_t(\omega)$ is a measurable function of (t, ω) on the product space $(\Gamma \times \Omega, \mathcal{B}(\Gamma) \otimes \mathcal{F}, \lambda \otimes P)$ where $\mathcal{B}(\Gamma)$ is the Borel σ-field on Γ and λ is the Lebesgue measure on Γ. In fact, if we let $X_{t,n}(\omega)$ to be defined as

$$X_{t,n}(\omega) = \sum_{k=1}^{n} Z_{k/n}(\omega) I_{[k/n,(k+1)/n]}(t),$$

we see that $Z_{t,n}(\omega)$ is a sum of products of functions which are $\mathcal{B}(\Gamma) \otimes \mathcal{F}$-measurable[2]. It follows that $Z_{t,n}(\omega)$ is measurable in (t, w), $n = 1, 2, \ldots$ and since $Z_t(\omega)$, by assumption, is continuous (right- or left-continuous) it can be written as a limit of measurable functions:

$$Z_t(\omega) = \lim_{n \to \infty} Z_{t,n}(\omega),$$

and it is thus $\mathcal{B}(\Gamma) \otimes \mathcal{F}$-measurable as well.

- $X_t(\omega)$ is integrable with respect to $\lambda \otimes P$ as by Tonelli's Theorem one has:

$$\begin{aligned}\int_{\Gamma \times \Omega} |X_t(\omega)|\, dP(\omega)\, dt &= \int_\Gamma \left(\int_\Omega |X_t(\omega)|\, dP(\omega) \right) dt \\ &= \int_\Gamma E[|Z_t|]\, dt \leq \left(\int_\Gamma E[X_t^2]\, dt \right)^{1/2} < \infty.\end{aligned}$$

[2]Assume that $\{Z_k : k \in \mathbf{Z}_+\}$ is a sequence of r.v.'s on $(\Omega, \mathcal{F}, P)$ and $\{Z_t^{(n)} : t \in [a, b]\}$ is a stochastic process defined as follows

$$Z_t^{(n)}(\omega) = Z_0(\omega) \cdot I_{[a,t_1)}(t) + \sum_{i=1}^{n-1} Z_{t_i}(\omega) \cdot I_{[t_i,t_{i+1})}(t) + Z_n(\omega) \cdot I_{[t_n,b]}(t)$$

for every partition $\{t_i : i = 0, 1, 2, \ldots, n-1, n, n+1\}$ such that $a = t_0 < t_1 < t_2 < \ldots < t_n < t_{n+1} = b$. Now, if $A \in \mathcal{B}(\mathbf{R})$, we find that

$$\begin{aligned}\{(t, \omega) : Z_t^{(n)}(\omega) \in A\} &= [a, t_1] \times \{\omega : Z_0(\omega) \in A\} \\ &\cup \left\{ \cup_{i=1}^{n-1} [t_i, t_{i+1}) \times \{\omega : X_{t_i}(\omega) \in A\} \right\} \cup [t_n, b] \times \{\omega : Z_n(\omega) \in A\}\end{aligned}$$

from which it follows easily that $\{(t, \omega) : Z_t^{(n)}(\omega) \in A\} \in \mathcal{B}([a, b]) \times \mathcal{F}$.

Since both λ and P are σ-finite measures (the assumption about Γ) we can apply Fubini's Theorem and conclude that $I_t(\omega) = \int_\Gamma X_t(\omega)\,dt$ is $\mathcal{F}$-measurable and, therefore, a random variable on $(\Omega, \mathcal{F})$.

Remark. If $\{X_t(\omega) : t \in \mathbf{R}_+\}$ is a real-valued stochastic process on a probability space $(\Omega, \mathcal{F}, P)$ and $\Gamma \subset \mathbf{R}_+$ is an interval, the very definition of stochastic process implies that $X_t(\cdot)$ is a $\mathcal{F}$-measurable funtion on Ω for every $t \in \Gamma$. From this, however, it does not follow that the mapping

$$X_\cdot(\cdot) : \Gamma \times \Omega \mapsto \mathbf{R}$$

is $\mathcal{B}(\Gamma) \otimes \mathcal{F}$-measurable as the next example shows.

Example. (Kallianpur (1980), p. 10) Let $X = \{X_t : t \in [0,1]\}$ be a stochastic process on a probability space $(\Omega, \mathcal{F}, P)$ consisting of mutually independent r.v.'s such that $E[X_t] = 0$ and $E[X_t^2] = 1$ for every $t \in [0,1]$. Then X is not measurable. In fact, if we assumed that it is measurable then there exists a (t,ω)-measurable family of stochastic functions $\{X_t(\omega)\}$ such that $E[X_t] = 0$, for $t \in [0,1]$, $E[X_s X_t] = 0$ if $s \neq t$ and $E[X_t X_s] = 1$ if $t = s$. Then, for every subinterval I of $[0,1]$ we have

Example of a non measurable process

$$\int_\Omega \int_I \int_I |X_s(\omega) X_t(\omega)|\, P(d\omega)\, ds\, dt < \infty.$$

Under these assumptions we are entitled to use the Fubini Theorem; hence

$$E\left[\left(\int_I X_t(\omega)\,dt\right)^2\right] = E\left[\int_I \int_I X_s(\omega) X_t(\omega)\,ds\,dt\right] = \int_I \int_I E[X_s X_t]\,ds\,dt = 0.$$

This is the same as saying that for a set A_I with $P(A_I) = 0$ we have $\int_I X_t(\omega)\,dt = 0$ if $\omega \notin A_I$. If we consider all subintervals $I = [r_i', r_i'']$ with $r_i', r_i'' \in \mathbf{Q}$ and let $A = \cup_I A_I$, we have established that $P(A) = 0$ and for all $\omega \in A^C$, $\int_a^b X_t(\omega)\,dt = 0$ for any subinterval $[a,b]$ of $[0,1]$. This means that $X_t(\omega) = 0$ for all t except possibly for a set of Lebesgue measure zero. Applying the Fubini Theorem again we find

$$\int_\Omega \int_0^1 X_t^2(\omega)\, P(d\omega)\, dt = 0$$

which is a contradiction as, by assumption, we know that

$$\int_\Omega \int_0^1 X_t^2(\omega)\, P(d\omega)\, dt = \int_0^1 E[X_t^2]\, dt = 1.$$

In establishing measurability of real-valued stochastic processes, the next lemma is sometimes useful.

Lemma. *Let $\{X_t : t \in \Gamma\}$ be a real-valued stochastic process and $\Gamma \subset \mathbf{R}_+$ a compact interval. If $X_t(\cdot)$ is continuous in probability, there exists a version of $X_t(\cdot)$ which is separable and measurable.*

Proof. See e.g. Todorovic (1992), Proposition 1.11.

Computing the distribution of the random variable I_Γ is not easy in general and as it depends on the distribution of the random variables X_t, $t \in \Gamma$.

Example. Let $\{W_t : t \in \mathbf{R}_+\}$ be Brownian motion with parameter σ. Computing the mean of $\int_0^1 W_t^2(\omega)\,dt$ is actually an easy task. In fact,

$$E[\int_0^1 W_t^2(\omega)\,dt] = \int_0^1 E[W_t^2(\omega)]dt = \int_0^1 \sigma^2 t dt = \sigma^2/2$$

since the existence of the integral above allows the interchanging of the operators E and $\int$. To compute the second moment, we need to use the fact that for the Wiener process the increments $W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$ are independent for $t_1 \le t_2 \le \dots \le t_{n-1} \le t_n$. Then,

$$\begin{aligned} E[\int_0^1 W_t^2\,dt]^2 &= E[\int_0^1 W_t^2\,dt \int_0^1 W_s^2\,ds] = \\ &= E[\int_0^1 (\int_0^1 W_t^2 W_s^2\,dt)\,ds] \end{aligned}$$

and, after interchanging E and $\int$,

$$\begin{aligned} &= \int_0^1 (\int_0^1 E[W_t^2 W_s^2]\,dt)\,ds \\ &= \int_0^1 (\int_0^s E[W_t^2 W_s^2]\,dt)\,ds + \int_0^1 (\int_s^1 E[W_t^2 W_s^2]\,dt)\,ds. \end{aligned}$$

When $t \le s$, one can write $W_s = [W_s - W_t + W_t]$ so that the first of the last two integrals above can be rewritten as

$$\begin{aligned} \int_0^1 (\int_0^s E[W_t^2 W_s^2]\,dt)\,ds &= \int_0^1 (\int_0^s E[W_t^2 (W_s - W_t + W_t)]^2\,dt)\,ds \\ &= \int_0^1 (\int_0^s E[W_t^2 (W_s - W_t)^2 + 2(W_s - W_t)W_t^3 + W_t^4]\,dt)\,ds \end{aligned}$$

and, using independence of $W_s - W_t$ and W_t in addition to basic properties of Wiener process,

$$= \sigma^4 \int_0^1 (\int_0^s [(s-t)t + 0 + 3t^2]dt)ds = \frac{7}{24}\sigma^4.$$

It is easily checked that the second of the two integrals above equals $(7/24)\sigma^4$ as well. In fact, since all integrals here exist, by Fubini's Theorem we can write

$$\int_0^1 (\int_s^1 E[W_t^2 W_s^2]\,dt)\,ds = \int_0^1 (\int_0^t E[W_s^2 W_t^2]\,ds)\,dt.$$

This implies that

$$E[\int_0^1 W_t^2\,dt]^2 = 2\sigma^4 \frac{7}{24} = \frac{7}{12}\sigma^4$$

and, therefore,

$$Var[\int_0^1 W_t^2\,dt] = \frac{7}{12}\sigma^4 - \left(\frac{\sigma^2}{2}\right)^2 = \frac{\sigma^4}{3}.$$

Problem A. Let $\{W_t : t \in \mathbf{R}_t\}$ be Brownian motion with parameter σ. Find the mean and variance of $\int_0^1 W_s\,ds$.

Solution. It is easily checked that $E[\int_0^1 W_s\,ds] = 0$ and that $E[[\int_0^1 W_s\,ds]^2 = \sigma^2/3$.

To complete this section on random integrals, we show how to compute the distribution of some relatively simple functionals. This is done working through an example that involves Brownian bridge.

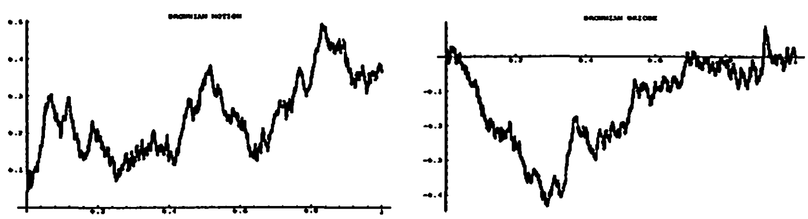

Figure 1.2: *Brownian motion ($\alpha = 0.0$, $\sigma = .3$) and its associated Brownian bridge.*

Example. Let $W^0 = \{W_t^0 : 0 \le t \le 1\}$ be a Brownian bridge and let $W = \{W_t : 0 \le t \le 1\}$ be the associated Wiener process. Define the simple linear functional

$$L = \int_0^1 g(t) W_t^0\,dt$$

where g is a continuously differentiable function on $[0,1]$ and $G(t) \equiv \int_0^t g(t)dt$, $0 \le t \le 1$, is square integrable. Then

$$L \sim N(0, A), \text{ where } A = \int_0^1 G^2(t)dt - (\int_0^1 G(t)dt)^2.$$

By definition, $W_t^0 = W_t - tW_1$, $0 \le t \le 1$, and it is easily proved that

$$(W_{t_1}^0, W_{t_2}^0, \dots W_{t_m}^0)' \sim N(\emptyset, \Gamma_m)$$

with

$$\Gamma_m = \begin{pmatrix} t_1 - t_1^2, & t_1 \wedge t_2 - t_1 t_2 & \dots & t_1 \wedge t_m - t_1 t_m \\ t_1 \wedge t_2 - t_1 t_2 & t_2 - t_2^2 & \dots & t_2 \wedge t_m - t_2 t_m \\ \dots & \dots & \dots & \dots \\ t_1 \wedge t_m - t_1 t_m & t_2 \wedge t_m - t_2 t_m & \dots & t_m - t_m^2. \end{pmatrix}$$

In addition, if we define

$$L_n = \sum_{i=0}^{n} g(i/n) W^0_{i/n} \frac{1}{n},$$

it is easy to check that

$$L = \lim_{n \to \infty} L_n$$

and

$$L_n \sim N(\emptyset, H'_n \Gamma_n H_n)$$

with $H'_n = (g(0), g(1/n), \dots, g((n-1)/n), g(1))$. The following lemma completes the proof of normality.

Lemma. Let X_n be a sequence of normal random variables converging in distribution to a random variable X. Then, X is either normal or constant.

Proof. If we look at the sequence of c.f.'s associated with X_n, we have

$$\phi_{X_n}(t) = \exp\{it\mu_n - (1/2)t^2\sigma_n^2\},$$

where both μ_n and $\sigma_n^2 \not\to \infty$ as $n \to \infty$ or otherwise there is not convergence in distribution as tightness is lost. If $\mu_n \to \mu$ and $\sigma_n^2 \to \sigma^2 < \infty$, then

$$\phi_{X_n}(t) \to \phi_X(t) = \begin{cases} \exp\{it\mu - (1/2)\sigma^2 t^2\} & \text{if } \sigma^2 \neq 0; \\ \exp\{it\mu\} & \text{if } \sigma^2 = 0. \end{cases}$$

This completes the proof of the lemma and proves the normality of the random functional.

To prove that the $E[L] = 0$ we use the fact that, under our assumption for g, the random variable $X_t(\omega) \equiv g(t) \cdot W^0_t(\omega)$ is clearly measurable and integrable. In addition, $X_t \sim N[0, g^2(t)(t - t^2)]$ and the result follows using Fubini's Theorem.

To compute the variance several integrations by parts are required. Assuming without loss of generality that $s < t$, we have

$$\begin{aligned} E[L^2] &= E[\int_0^1 \int_0^1 g(s)g(t) W^0_s W^0_t \, dsdt] \\ &= 2E[\int_0^1 \int_s^1 g(s)g(t) W^0_s W^0_t \, dsdt] \\ &= 2 \int_0^1 \int_s^1 g(s)g(t) E[W^0_s W^0_t] \, dsdt \\ &= 2 \int_0^1 \int_s^1 g(s)g(t)s(1-t) dsdt \end{aligned}$$

using again Fubini's Theorem and the fact that for Brownian bridge

$$E[W^0_s W^0_t] = s \wedge t - st.$$

Letting $u = 1-t$ and $dv = g(t)dt$, we have $du = -dt$ and $v = G(t)$ and, therefore, we get

$$\begin{aligned}E[L^2] &= 2\int_0^1 sg(s)ds \int_s^1 g(t)(1-t)dt = \int_0^1 sg(s)ds\big[-G(t)(1-t)\mid_s^1 \\ &+ \int_s^1 G(t)dt\big] = \int_0^1 sg(s)\big[-G(s)(1-s) + \int_s^1 G(t)dt\big]ds \\ &= \int_0^1 -2s(1-s)g(s)G(s)ds + 2sg(s)\big[\int_s^1 G(t)dt\big]ds.\end{aligned}$$

Now, let $u = -s(1-s)$, $dv = 2g(s)G(s)ds$ so that $du = (-1+2s)ds$ and $v = G^2(s)$. Then,

$$\int_0^1 -2s(1-s)g(s)G(s)ds = -2s(1-s)G^2(s)\mid_0^1 + \int_0^1 G^2(s)ds - \int_0^1 2sG^2(s)ds.$$

In addition, it is easily checked that

$$-2s(1-s)G^2(s)\mid_0^1 = 0$$

so that it is possible to write

$$E[L^2] = \int_0^1 G^2(s)ds - \int_0^1 2sG^2(s)ds + 2\int_0^1 sg(s)\big[\int_s^1 G(t)dt\big]ds.$$

The last partial integration involves $2\int_0^1 sg(s)\big[\int_s^1 G(t)dt\big]ds$. To this purpose, let $u = s$ and $dv = g(s)ds$. Then, $du = ds$ and $v = G(s)$. Thus, using Fubini's Theorem, we have

$$\begin{aligned}\int_0^1 2sg(s)\big[\int_s^1 G(t)dt\big]ds &= \int_0^1\int_0^t g(s)sG(t)dsdt \\ &= 2\big[\int_0^1 G(s)s\mid_0^t - \int_0^t G(s)ds\big]G(t)dt = 2\int_0^1 \big[G(t)t - \int_0^t G(s)ds\big]G(t)dt \\ &= 2\int_0^1 G^2(t)tdt - 2\int_0^1\int_0^t G(s)G(t)dsdt = 2\int_0^1 G^2(t)tdt - \big[\int_0^1 G(t)dt\big]^2.\end{aligned}$$

Thus,

$$\begin{aligned}E[L^2] &= \int_0^1 G^2(s)ds - \int_0^1 2sG^2(s)ds + \int_0^1 2G^2(t)tdt - \big[\int_0^1 G(t)dt\big]^2 \\ &= \int_0^1 G^2(s)ds - \big[\int_0^1 G(t)dt\big]^2\end{aligned}$$

as we were supposed to prove. In particular, because of the assumptions that G is square integrable, the variance of the limit distribution is finite.

Problem B. Let $W = \{W_t : 0 \leq t \leq 1\}$ be Brownian motion and let $H \equiv \int_0^1 g(t)W_t(\omega)\,dt$. Prove that

$$H \sim N(0, B) \text{ where } B = \int_0^1 G^2(t)dt + G^2(1) - 2G(1)\int_0^1 G(t)dt,$$

where g is a continuously differentiable function on $[0,1]$ and $G(t) \equiv \int_0^t g(t)dt$. $0 \leq t \leq 1$, is square integrable.

Solution. Since using the definition of Brownian bridge it is easy to show that

$$H = L + W_1 \int_0^1 tg(t)\,dt,$$

and we proved that $L = \int_0^1 g(t)W_t^0(\omega)\,dt$ is normal, H, is a linear combination of normal r.v.'s and, thus, it is also normally distributed.

To simplify the computations of the moments of the distribution of H we can use the following fact which is easily established using integration by parts.

$$\int_0^1 tg(t)\,dt = G(1) - \int_0^1 G(t)\,dt.$$

Using this result it is easy to prove that $E[H] = 0$. To this purpose it suffices to write

$$E[H] = E[L] + E[W_1][G(1) - \int_0^1 G(t)dt] = 0 + 0[G(1) - \int_0^1 G(t)dt] = 0.$$

Clearly, we have also

$$\begin{aligned}\mathrm{Var}[H] = \mathrm{Var}[L] + \mathrm{Var}[W_1G(1) - W_1\int_0^1 G(t)dt] + \\ 2\mathrm{Cov}[L, W_1G(1) - W_1)\int_0^1 G(t)dt].\end{aligned}$$

$\mathrm{Var}[L]$ was computed above. In addition, as $E[W(1)G(1) - W(1)\int_0^1 G(t)dt] = 0$, we have that $\mathrm{Var}[W_1G_1 - W(1)\int_0^1 G(t)dt] = E[W_1G(1) - W_1\int_0^1 G(t)dt]^2$. Hence,

$$\begin{aligned}E[W_1G(1) - W_1\int_0^1 G(t)dt]^2 &= E[W_1^2G^2(1) + W_1^2\Big[\int_0^1 G(t)dt\Big]^2 \\ -2W_1^2G(1)\int_0^1 G(t)dt] &= 1 \cdot G^2(1) + \Big[\int_0^1 G(t)dt\Big]^2 - 2G(1)\int_0^1 G(t)dt.\end{aligned}$$

Finally,

$$\begin{aligned}&\mathrm{Cov}[L, W(1)G(1) - W(1)\int_0^1 G(t)dt] \\ &\quad = E[L \cdot W(1)[G(1) - \int_0^1 G(t)dt]] = [G(1) - \int_0^1 G(t)dt]E[LW(1)]\end{aligned}$$

and, with simple manipulations, we find

$$E[LW_1] = E[\int_0^1 g(t)W_tW_1dt - \int_0^1 g(t)tW_1^2dt]$$
$$= \int_0^1 g(t)E[W_tW_1]dt - \int_0^1 tg(t)E[W_1^2]dt$$
$$= \int_0^1 g(t)E[W_t^2]dt - \int_0^1 tg(t)E[W_1^2]dt = \int_0^1 tg(t)dt - \int_0^1 tg(t)\cdot 1dt = 0$$

where the fact that for a Wiener process $E[W_tW_s] = E[W_s^2] = s$, $s \leq t$, was used.

3. Itô Stochastic Integral

For certain M and X, the integral can be defined path-by-path. For instance, if M is a right-continuous local L^2-martingale whose paths are locally of bounded variation, and X is a continuous adapted process, then $\int_{[0,t]} X_s(\omega)dM_s(\omega)$ is well-defined as a Riemann-Stieltjes integral for each t and ω, namely by the limit

$$\lim_{n\to\infty} \sum_{k=0}^{[2^nt]} X_{k2^{-n}}(\omega)\cdot(M_{(k+1)2^{-n}}(\omega) - M_{k2^{-n}}(\omega)).$$

This definition works for the case in which $M_t = N_t - \lambda t$, i.e. M where N is a Poisson process of parameter λ. If the process X is bounded, adapted, and has continuous sample paths, one can prove that

$$I_t = \int_0^t X_s dM_s = \sum_{n=1}^{\infty} X_{T_n} I_{\{t\geq T_n\}} - \lambda \int_0^t X_s ds$$

where $\{T_n : n \in \mathbb{N}\}$ are the arrival times of the Poisson process N.
Unfortunately, the path-by-path definition doesn't work all the time. In fact, one can show that the limit above exists for every continuous stochastic process X if and only if M is a finite variation process, i.e. almost all its paths are of finite variation on each compact interval of $\mathbb{R}_+$ (see e.g. (Protter, 1990; Theorem 1.49). Essentially, a stochastic process A is of finite variation if for any $[s,t]$ $s,t \in \mathbb{R}_+$, there exists a constant K such that for every partition $\{s = t_0 < t_1 < \ldots < t_n = t\}$

finite variation process

$$\sum_{k=0}^{n-1} |A_{t_{k+1}} - A_{t_k}| \leq K.$$

There are interesting examples of stochastic processes that do not satisfy this requirement, the most important being provided by Brownian motion.

Example. A Browian motion has infinite variation on any interval, no matter how small it is. This is the consequence of two facts:

(1) If g is continuous and of finite variation, then its quadratic variation is zero (see e.g. Klebaner 1998; Theorem 1.10 for a proof);

(2) The quadratic variation of a Brownian motion over $[0,t]$ is t. (see e.g. Klebaner 1998; Theorem 3.5 for a proof).

Thus, the stochastic integral $\int_{[0,t]} W dW$ where W is a Brownian motion in $\mathbf{R}$ would not be defined. This negative result holds not only for the case of Brownian motion but for any continuous martingale M.

Example. Assume that $M_0 = 0$, M has finite variation then, introducing the partition $\Delta_n = \{0 = t_0 < t_1 < \ldots < t_n = t\}$, it is possible to write

$$\begin{aligned} E[M_t^2] &= E[\sum_{i=0}^{n-1} M_{t_{i+1}}^2 - M_{t_i}^2] \\ &= E[\sum_{i=0}^{n-1} (M_{t_{i+1}} - M_{t_i})^2] \end{aligned}$$

using the fact that M is a martingale and, therefore, $M_{t_{i+1}} \cdot M_{t_i} = E[E[M_{t_{i+1}} \cdot M_{t_i} \mid \mathcal{F}_{t_i}]] = M_{t_i} \cdot E[M_{t_{i+1}} \mid \mathcal{F}_{t_i}] = M_{t_i}^2$

$$\begin{aligned} &\leq E[\sup_{0\leq i\leq n-1} |M_{t_{i+1}} - M_{t_i}| \cdot \sum_{i=0}^{n-1} |M_{t_{i+1}} - M_{t_i}|] \\ &\leq E[\sup_{0\leq i\leq n-1} |M_{t_{i+1}} - M_{t_i}| \cdot K] \end{aligned}$$

and letting $n \to \infty$ so that $\max_{0\leq i\leq n-1} t_{i+1} - t_i \to 0$ we have that $|M_{t_{i+1}} - M_{t_i}| \to 0$ as, by assumption, M is a continuous martingale. This proves that $E[M_t^2] = 0$ and, hence, $M_t = 0$ a.s. Essentially this says that the only continuous martingale with respect to which a path-by-path integration is defined, is a stochastic process which is a.s. equal to 0 for every $t \in \mathbf{R}_+$.

A different and more general definition of stochastic integral is called for. Most of this chapter is devoted to defining the stochastic integral $\int_{[0,t]} X dW$ known as Itô integral. This integral is not defined path-by-path but via an isometry between a space of processes that are square integrable with respect to a measure induced by M, and a space of square integrable stochastic integrals $\int X dM$.

Itô integral

Suppose $0 \leq S < T$, $W_t(\omega)$ is a 1-dimensional Brownian motion starting at the origin and let $f : \mathbf{R}_+ \times \Omega \mapsto \mathbf{R}$ be a given function. We want to define

$$\int_S^T f(t,\omega) dW_t(\omega). \tag{1.1}$$

It is reasonable to start with a definition for a simple class of functions f and then extend it by some approximation procedure. Thus, let us assume that f

has the form

$$f(t,\omega) = \sum_{j=0}^{[2^n t]-1} c_j(\omega) I_{[j2^{-n},(j+1)2^{-n})}(t), \ n \in \mathbf{N}. \tag{1.2}$$

For such functions it is reasonable to define

$$\int_S^T f(t,\omega)\, dW_t(\omega) = \sum_{j=0}^{[2^n t]-1} c_j(\omega)[W_{t_j^{(n)}+1}(\omega) - W_{t_j^{(n)}}(\omega)]$$

where

$$t_j^{(n)} = \begin{cases} k2^{-n} & \text{if } S \le k2^{-n} \le T; \\ S & \text{if } k2^{-n} < S; \\ T & \text{if } k2^{-n} > T. \end{cases}$$

However, without any further assumptions on the functions $c_j(\omega)$ this leads to difficulties as the next example shows.

Example. Let

$$\begin{aligned} f_1(t,\omega) &= \sum_{j=0}^{[2^n t]-1} W_{j2^{-n}}(\omega) \cdot I_{[j2^{-n},(j+1)2^{-n})}(t) \\ f_2(t,\omega) &= \sum_{j=0}^{[2^n t]-1} W_{(j+1)2^{-n}}(\omega) \cdot I_{[j2^{-n},(j+1)2^{-n})}(t) \end{aligned} \tag{1.3}$$

then,

$$E[\int_0^T f_1(t,\omega)\, dW_t(\omega)] = \sum_{j=0}^{[2^n t]-1} E[W_{t_j^{(n)}}(\omega)] \cdot E[(W_{t_j^{(n)}+1}(\omega) - W_{t_j^{(n)}}(\omega)] = 0$$

using the fact that for a Brownian motion $W_{t_{j(n)}+1}(\omega) - W_{t_j^{(n)}}(\omega)$ is independent of $W_{t_j^{(n)}}(\omega)$ for all $t_j^{(n)}$, $j = 0, 1, 2, \dots, [2^n t] - 1$. (Here and in the following of this chapter E is the expectation with respect to the law P^0 for a Brownian motion starting at 0). But,

$$E[\int_0^T f_2(t,\omega)\,dW_t(\omega)] = E[\sum_{j=0}^{[2^n t]-1} W_{t_{j(n)+1}}(\omega)\cdot[(W_{t_{j(n)+1}}(\omega) - W_{t_j^{(n)}}(\omega)]$$
$$= \sum_{j=0}^{[2^n t]-1} E[W^2_{t_{j(n)+1}}(\omega)] + E[W_{t_{j(n)+1}} \cdot W_{t_j^{(n)}}(\omega)]$$
$$= \sum_{j=0}^{[2^n t]-1} t_{j(n)+1} - \sum_{j=0}^{[2^n t]-1} E[W^2_{t_j^{(n)}}(\omega)]$$
$$+ \sum_{j=0}^{[2^n t]-1} E[(W_{t_{j(n)+1}}(\omega) - W_{t_j^{(n)}}(\omega))\cdot W_{t_j^{(n)}}(\omega)]$$
$$= \sum_{j=0}^{[2^n t]-1} (t_{j(n)+1} - t_{j(n)}) + 0 = T.$$

So, in spite of the fact that both $f_1(\cdot)$ and $f_2(\cdot)$ appear to be very reasonable approximation to $f(t,\omega) = W_t(\omega)$, their integrals according to (1.1) are not close to each other at all, no matter how large n is chosen. This reflects the fact that variations of the paths of W_t are too big to enable us to define $\int_S^T f(t,\omega)dW_t(\omega)$ in the Riemann-Stielties sense. In fact, one can show that the paths $t \mapsto W_t$ of a Brownian motion are nowhere differentiable almost everywhere.

In general, it is natural to approximate a given function $f(t,\omega)$ by

$$\sum_j f(t_j^*,\omega) I_{[t_j,t_j+1)}(t)$$

where $t_j^* \in [t_j, t_j+1]$, and then define $\int_S^T f(t,\omega)dW_t(\omega)$ as the limit (in some sense to be explained later) of $\sum_j f(t_j^*,\omega)\cdot[(W_{t_j^{(n)}+1}(\omega) - W_{t_j^{(n)}}(\omega)]$ as $n \to \infty$. However, as the example above shows, unlike the Riemann-Stielties integral, it does make a difference what points t_j^* are chosen. The following two choices have turned out to be the most useful ones:

(i) $t_j^* = t_j$ (left-end point) which leads to the **Itô integral**;

(ii) $t_j^* = (t_j + t_{j+1})/2$ (mid-point) which leads to the **Stratonovich integral**.

In any case, one must restrict oneself to a special class of functions $f(t,\omega)$ also if they have the form (1.2) in order to obtain a reasonable definition of the integral. We will work with Itô's choice, $t_j^* = t_j$. The approximation procedure will work out fine provided that $f()$ has the property that each of the functions $\omega \mapsto f(t_j,\omega)$ only depends on the behavior of $W_s(\omega)$ up to time t_j. This leads to the next definition.

Definition 1. *Let $\{X_t(\omega) : t \in \mathbf{R}_+\}$ be a stochastic process from a probability space $(\Omega, \mathcal{F}, P)$ into $\mathbf{R}^n$. Then, define $\mathcal{F}_t$ to be the σ-algebra generated by the random variables $X_s(\omega), s \leq t$. In other words, $\mathcal{F}_t$ is the smallest σ-algebra containing all sets of the form*

$$\{\omega : X_{t_1}(\omega) \in F_1, \dots , X_{t_k}(\omega) \in F_k\}$$

$F_k \in \mathcal{B}(\mathbf{R}^n)$, $k = 1, 2, \dots$, $F_j \in \mathcal{B}(\mathbf{R}^n)$. *In addition, we assume that all sets of measure zero are included in $\mathcal{F}_t$.*

Theorem 2. (Exercise 3.14 in Øksendal (1995)) *A function $h(\omega)$ defined on $(\Omega, \mathcal{F}, P)$ will be $\mathcal{F}_t$-measurable iff it can be written as the pointwise a.e.[P] limit of sums of functions of the form*

$$g_1(X_{t_1}(\omega)) \cdot g_2(X_{t_2}(\omega)) \cdot \dots \cdot g_k(X_{t_k}(\omega))$$

where $g_1(\cdot), g_2(\cdot), \dots , g_k(\cdot)$ are continuous functions and $t_j \leq t$ for $j \leq k$, $k = 1, 2, \dots$

Proof. For $n, j = 1, 2, \dots$ put $t_j = t_j^{(n)} = j \cdot 2^{-n}$. For fixed n, let $\mathcal{H}_n$ be the σ-algebra generated by $\{X_{t_j}(\cdot), t_j \leq t\}$. Then, clearly $\mathcal{H}_n \uparrow \mathcal{F}_t$ as $n \to \infty$. Then, we need the following lemma (see e.g. Øksendal 1995, p. 243 for a proof),

Lemma A. *Let $X \in L^1(P)$, and let $\{N_k, k \in \mathbf{N}\}$ be an increasing family of σ-algebras such that $\mathcal{N}_k \subset \mathcal{F}$ $k = 1, 2, \dots$ and define $\mathcal{N}_\infty$ to be the σ-algebra $\sigma\left(\cup_{k=1}^{\infty} \mathcal{N}_k\right)$. Then,*

$$E[X \mid \mathcal{N}_k] \to E[X \mid \mathcal{N}_\infty] \text{ as } k \to \infty$$

a.e.[P] and in $L^1(P)$.

If we let $h_n \equiv E[h \mid \mathcal{H}_n]$, since $h = E[h \mid \mathcal{F}_t]$ as by assumption $h()$ is $\mathcal{F}_t$-measurable we can use Lemma A and state that

$$h = E[h \mid \mathcal{F}_t] = \lim_n E[h \mid \mathcal{H}_n] = \lim_n h_n \quad \text{a.e}[P] \text{ pointwise.}$$

For the next step we need a second lemma (see. e.g. Breiman, Probability; 1968, Proposition A.21 p. 395 for a proof)

Lemma B. *If $X, Y : \Omega \mapsto \mathbf{R}^n$ are two given functions, then Y is $\mathcal{H}^X$-measurable function iff there exists a Borel measurable funtion $g : \mathbf{R}^n \mapsto \mathbf{R}^n$ such that $Y = g(X)$.*

Since h_n is $\mathcal{H}_n$-measurable because of the way it was defined, one can use Lemma B to state that it must be

$$h_n(\omega) = G_n(X_{t_1}(\omega), X_{t_2}(\omega), \dots X_{t_k}(\omega))$$

for some Borel function $G_n : \mathbf{R}^k \mapsto \mathbf{R}$, where $k = \max\{j : j \cdot 2^{-n} \leq t\}$. The third step requires the following (see e.g. Nathanson, Theory of Functions of Real Variable, Vol I (1964) Theorem 4, p. 114 for a proof)

Lemma C. *Every Borel-measurable function $G : \mathbf{R}^k \mapsto \mathbf{R}$ which is finite a.e.[P] can be approximated pointwise a.e.[P] by a continuous function $F : \mathbf{R}^k \mapsto \mathbf{R}$.*

This last Lemma allows us to state that h_n can be approximated by a continuous funtion F_n from $\mathbf{R}^k \mapsto \mathbf{R}$.
The next stage is a little bit more difficult and it is based on the Stone-Weierstrass Theorem. Since we want our theorem to be general enough to handle as many cases as possible, we will assume that the stochastic process $\{X_t(\omega) : t \in \mathbf{R}\}$ is an $\mathbf{R}$-valued r.v. for each $t \in \mathbf{R}_+$. In this way we need to use a version of the Stone-Weierstrass Theorem that fits the case where the spaces involved are not compact (see Folland 1986, Theorem 4.52).

Theorem (Stone-Weierstrass). *Let X be a noncompact, locally compact Hausdorff space. If $\mathcal{A}$ is a closed subalgebra of $C_0(X, \mathbf{R}) = C_0(X) \cap C(X, \mathbf{R})$ (where $C(X, \mathbf{R})$ is the collection of continuous function from X into $\mathbf{R}$ and $C_0(X) = \{f \in C(X) :$ vanishes at $\infty\}$) which separates points, then either $\bar{\mathcal{A}} = C_0(X, \mathbf{R})$ or $\bar{\mathcal{A}} = \{f \in C_0(X, \mathbf{R}); f(x_0) = 0\}$ for some $x_0 \in X$. The former case holds when $\mathcal{A}$ contains the constants.*

Now, $\mathbf{R}^k$ is clearly a noncompact, locally compact Hausdorff space. So, if we can find the appropriate algebra, we are done. To this purpose, we use the following

Lemma D. *Let $X_1, X_2, \dots X_k$ be locally compact metric space. Consider the Cartesian product $\Pi_{i=1}^k X_i$. Then*[3] *$\Pi_{i=1}^k X_i$ is locally compact. In addition, if $f \in C(\Pi_{i=1}^k X_i)$ and $\epsilon > 0$, then there exist functions $\{f_{i1}, f_{i2}, \dots f_{im}\} \in C(X_i)$, $i = 1, 2, \dots, k$ such that*

$$|f(x_1, \dots x_k) - \sum_{i=1}^{m} \Pi_{j=1}^k f_{ij}(x_j)| < \epsilon$$

holds for all $(x_1, x_2, \dots, x_k) \in \Pi_{i=1}^k X_i$.

Proof. Let

$$\mathcal{A} = \{f \in C(\Pi_{i=1}^k X_i) : \exists (f_{i1}, \dots f_{im}) \in C(X_i), i = 1, 2, \dots, k$$
$$\text{and } f(x_1, \dots x_n) = \sum_{i=1}^{m} \Pi_{j=1}^k f_{ij}(x_j)\ \forall (x_1, \dots x_k) \in \Pi_{i=1}^k X_i\}.$$

It is simple to verify that $\mathcal{A}$ is an algebra of functions of $C(\Pi_{i=1}^k X_i)$ and that the constants are in $\mathcal{A}$. In addition, if $(x_1, \dots x_k) \neq (y_1, \dots, y_k)$ then it must be $x_i \neq y_i$ for at least one i, $i = 1, 2, \dots, k$. Assume that it is $x_j \neq y_j$. In this case, let $f(x_1, \dots, x_k) = f_j(x_j)$ for some $f_j \in C(X_j)$ such that $f_j(x_j) \neq f_j(y_j)$. The same argument can be repeated for each of the other indexes. In any case, $f \in \mathcal{A}$ (note that the constants are in $\mathcal{A}$) and $f(x_1, \dots x_k) \neq f(y_1, \dots, y_k)$. Hence $\mathcal{A}$ separates the points of $\Pi_{i=1}^k X_i$ and hence, by the Stone-Weierstrass Theorem we have that $\bar{\mathcal{A}} = C_0(\Pi_{i=1}^k X_i, \mathbf{R})$.

If we replace $\Pi_{i=1}^k X_i$ with $\mathbf{R}^k$ we find that every h_n can be approximated by means of $\sum_{i=1}^m \Pi_{j=1}^k f_{ij}(x_j)$.

[3]Use the following fact (see e.g. J. Dugundji, Topology; 1966, Theorem 6.2.(4)): $\Pi_{\alpha \in A} Y_\alpha$ is locally compact iff all the Y_α are locally compact and at most finitely many are not compact. In our case where A is finite, the second part of the statement is certainly true.

To show that this holds for the limit function h as well consider that for any choice of $\epsilon > 0$ there exists $\bar{n}_\epsilon$ such that $|h(\cdot) - h_n(\cdot)| < \epsilon/3$ for all $n > \bar{n}_\epsilon$ and $|h_n(\cdot) - h_m(\cdot)| < \epsilon/3$, $\forall n, m > \bar{n}_\epsilon$. So, fix $n^* > \bar{n}_\epsilon$; because of Lemma D, we know that there exist k^* and $p(\epsilon, n^*)$ such that

$$|h_{n^*}(x_1, \ldots x_{k^*}) - \sum_{i=1}^{p} \Pi_{j=1}^{k^*} f_{ij}(x_j)| < \epsilon/3$$

for all $p > p(\epsilon, n^*)$. So, if we call $\sum_{i=1}^{p} \Pi_{j=1}^{k^*} f_{ij}(x_j) \equiv \psi$, we find that

$$|h - \psi| \leq |h - h_n| + |h_n - h_{n^*}| + |h_{n^*} - \psi| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

for all $n > \bar{n}_\epsilon$. The reverse implication clearly holds and the proof is complete.

The theorem above is handy in some situations. For instance, it can be used to state that $h(\omega) = W_{t/2}(\omega)$ is $\mathcal{F}_t$-adapted, while $h(\omega) = W_{2t}(\omega)$ is not. Nonetheless, when dealing with Brownian motion this result is not really needed as it is possible to exploit some of the features of this particular stochastic process. In fact, one can prove that $h(\omega) = W_{2t}(\omega)$ is not $\mathcal{F}_t$-adapted using the fact that $W_{2t}(\omega) = [W_{2t}(\omega) - W_t(\omega)] + W_t(\omega)$ and $W_{2t}(\omega) - W_t(\omega)$ is independent of $\mathcal{F}_t$. Now, if it were $\sigma(W_{2t} - W_t) \subset \mathcal{F}_t$, then every $A \in \sigma(W_{2t} - W_t)$ should have measure 0 or 1 which is not true. Hence there must be at least one event A in $\sigma(W_{2t} - W_t)$ which is not in $\mathcal{F}_t$, thus proving that $h(\omega) = W_{2t}(\omega)$ is not $\mathcal{F}_t$-adapted.

Definition 3. *Let $\{\mathcal{N}_t : t \in \mathbf{R}_+\}$ be an increasing families of σ-algebras of subsets of Ω. A function $g(t, \omega) : \mathbf{R}_+ \times \Omega \mapsto \mathbf{R}$ is called $\mathcal{N}_t$-adapted if for all $t \geq 0$ the function $\omega \mapsto g(t, \omega)$ is $\mathcal{N}_t$-measurable.*

Definition 4. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, $W = \{W_t : t \in \mathbf{R}_+\}$ a Brownian motion, and $\{\mathcal{F}_t : t \in \mathbf{R}_+\}$ a filtration of sub-σ-algebras of $\mathcal{F}$ such that W_t is $\mathcal{F}_t$-measurable with $E[W_t \mid \mathcal{F}_0] = 0$ and $E[W_t - W_s \mid \mathcal{F}_s] = 0$ w.p.1, for all $0 \leq s \leq t$. For $0 < T < \infty$ we define the class $\mathcal{L}_T^2$ as the class of functions* $\mathcal{L}_T^2$-class

$$f : [0, T] \times \Omega \mapsto \mathbf{R}$$

such that[4]

(a) f is $\mathcal{B}_{[0,T]} \times \mathcal{F}$-measurable, where $\mathcal{B}_{[0,T]}$ denotes the Borel σ-algebra on $[0, T]$;

[4] Condition (c) is not superfluous. In fact, for any $1/4 < t_0 < T$, let

$$f(t, \omega) = \begin{cases} W_t(\omega) & \text{if } t \in [0, T] - \{t_0\} \\ \exp\{W_{t_0}^2(\omega)\} & \text{if } t = t_0. \end{cases}$$

In this case, $E[f(t, \omega)^2] = \infty$ when $t = t_0$ but

$$\int E[f(t, \cdot)^2]\, dt < \infty$$

as $\{t_0\}$ is a zero-measure set.

(b) $\int_0^T E[f(t,\cdot)^2]dt < \infty$;

(c) $E[f(t,\cdot)^2] < \infty$ *for each* $0 \le t \le T$;

(d) $f(t,\cdot)$ *is* $\mathcal{F}_t$*-measurable for each* $0 \le t \le T$.

In addition, we consider two functions in $\mathcal{L}_T^2$ *to be identical if they are equal for all* (t,ω) *except possibly on a subset of* $\lambda \times P$*-measure zero. Then, with the norm*

$$\| f \|_{2,T} \equiv \sqrt{\int_0^T E[f(t,\cdot)^2]dt}$$

$\mathcal{L}_T^2$ *is a complete normed space, i.e. a Banach space, provided we identify functions which differ only on sets of measure zero.*

Note. Conditions (a)-(d) are stronger than requiring $f \in L^2([0,T]\times\Omega, \mathcal{B}_{[0,T]} \times \mathcal{F}, \lambda \times P)$ which by Fubini's Theorem guarantees (c) for $a.a.[\lambda]$ $t \in [0,T]$.

$\mathcal{S}_T^2$–class

Denote by $\mathcal{S}_T^2$ be subset of all step functions in $\mathcal{L}_T^2$. The next theorem states that we can approximate any function in $\mathcal{L}_T^2$ by step functions in $\mathcal{S}_T^2$ to any desired degree of accuracy in the norm $\|\cdot\|_{2,T}$.

Theorem 5. $\mathcal{S}_T^2$ *is dense in* $(\mathcal{L}_T^2, \|\cdot\|_{2,T})$.

Proof. Let's start considering partitions of $[0,T]$ of the form $0 = t_1^{(n)} < t_2^{(n)} < \ldots < t_{n+1}^{(n)} = T$ with $t_{i+1}^{(n)} - t_i^{(n)} \to 0$ for $i = 1,2,\ldots,n$ as $n \to \infty$. For any partition $0 = t_1 < t_2 < \ldots < t_{n+1} = T$ and any mean-square integrable $\mathcal{F}_{t_j}$-measurable random variables $f_j(\omega)$, $j = 1,2,\ldots,n$ we define a step function $\phi \in \mathcal{L}_T^2$ by

$$\phi(t,\omega) = \sum_{k=1}^{n} f_j(\omega) I_{(t_j,t_{j+1}]}(t), \;\; j = 1,2,\ldots,n.$$

A sequence of step functions $\{\phi_n, n \in \mathbf{N}\}$ in $\mathcal{L}_T^2$ can then be defined by

$$\phi_n(t,\omega) = \sum_{k=1}^{n} f_j^{(n)}(\omega) I_{(t_j^{(n)},t_{j+1}^{(n)}]}(t) \;\; w.p.1.$$

Clearly, $\phi_n(\cdot,\cdot) \in \mathcal{S}_T^2$ for each $n = 1,2,\ldots$

bounded continuous functions

Lemma 1. *Let* $g \in \mathcal{L}_T^2$ *be bounded and continuous for each* $\omega \in \Omega$*. Then, there exist step functions* $\phi_n \in \mathcal{S}_T^2$ *such that*

$$\int_0^T E\Big(|g(t,\cdot) - \phi_n(t,\cdot)|^2\Big)dt \to 0 \text{ as } n \to \infty.$$

Proof. Define $\phi_n(t,\omega)$ by

$$\phi_n(t,\omega) = \sum_{k=1}^{n} g(t_j^{(n)},\omega) I_{(t_j^{(n)},t_{j+1}^{(n)}]}(t).$$

Then, $\phi_n(\cdot,\cdot) \in \mathcal{S}_T^2$ since $g(t_j^{(n)},\omega)$ is $\mathcal{F}_{t_j^{(n)}}$-measurable, and when $n \to \infty$, we have $\phi(t,\omega) \to g(t,\omega)$ *a.a.*$[P]\,\omega$ and

$$E\Big(|g(t,\cdot) - \phi_n(t,\cdot)|^2\Big) \to 0,$$

for each $t \in [0,T]$, using the continuity of g. Now, using the boundedness part of the assumption about g, we can use the Bounded Convergence Theorem applied to the space of functions $(L^1[0,T], \mathcal{B}_{[0,T]}, \lambda)$ and state that

$$\int_0^T E\Big(|g(t,\cdot) - \phi_n(t,\cdot)|^2\Big)dt \to 0$$

as $n \to \infty$ as we were supposed to prove.

Lemma 2. *Let $h \in \mathcal{L}_T^2$ be bounded. Then there exist bounded functions $g_n \in \mathcal{L}_T^2$ such that $g_n(\cdot,\omega)$ is continuous for all ω and n, and* bounded functions

$$\int_0^T E\Big(|h(t,\cdot) - g_n(t,\cdot)|^2\Big)dt \to 0 \text{ as } n \to \infty.$$

Proof. Suppose $|h(t,\omega)| \le M < \infty$ for all (t,ω). For each n let ψ_n be a non-negative continuous function on $\mathbf{R}$ such that $\psi(x) = 0$ for $x \le -1/n$ and $x \ge 0$ and $\int_{\mathbf{R}} \psi_n(x)dx = 1$ for all $n = 1,2,\ldots$. Such a function is, for instance, provided by

$$\psi_n(x) = \begin{cases} -4n^2x & \text{if } -1/(2n) \le x \le 0; \\ 4n^2x + 4n & \text{if } -1/n \le x < -1/(2n); \\ 0 & \text{otherwise.} \end{cases}$$

Define

$$g_n(t,\omega) = \int_0^t \psi_n(s-t)h(s,\omega)ds, \quad n = 1,2,\ldots$$

The continuity of the g_n's for each ω is easy to check as

$$|g_n(t+h,\omega) - g_n(t,\omega)| \le \int_0^t |\psi_n(s-t-h,\omega) - \psi_n(s-t,\omega)|\cdot|h(s,\omega)|ds$$
$$+ \int_t^{t+h} \psi_n(s-t-h)\cdot|h(s,\omega)|ds,$$

$h()$ is bounded, and $\psi_n()$ is a continuous function, for each $n = 1,2,\ldots$ The $\mathcal{F}_t$-measurability of $g_n(t,\cdot)$ follows from the fact that $h \in \mathcal{L}_T^2$ and the integral can be regarded as the limit of sums. Now, we need to use the following result about approximate identities (see e.g. Jones (1993); p. 285-6 for a proof.)

Theorem. *Let $\{\phi_k, k \in \mathbf{N}\}$ be a sequence of functions in $L^1(\mathbf{R})$ such that*

- $\lim_k \int \phi_k(x)dx = c$ *exists;*
- $\int |\phi_k(x)|dx \le M < \infty$;
- $\lim_k \int_{|x|>r} |\phi_k(x)|dx = 0$ *for all* $r > 0$.

Then, for every $f \in L^p(\mathbf{R}^n)$, $1 \le p < \infty$, *we have*

$$\lim_k \| f \star \phi_k - cf \|_p = 0.$$

This theorem holds in our case replacing ϕ_k with ψ_n, c with 1, and $p = 2$. Hence, we have established that

$$\int_0^T [g_n(s,\omega) - h(s,\omega)]^2 ds \to 0$$

as $n \to \infty$ for each ω. So, by bounded convergence, we can state that

$$E\left(\int_0^T [g_n(t,\omega) - h(t,\omega)]^2 dt\right) \to 0 \text{ as } n \to \infty.$$

Since all integrals (involving bounded functions only) exists, we can use Fubini's Theorem and interchange the order of integration. This gives

$$\int_0^T E\left(h(t,\omega) - g_n(t,\omega)\right)^2 dt \to 0 \text{ as } n \to \infty.$$

$\mathcal{L}_T^2$ **Lemma 3.** *Let* $f \in \mathcal{L}_T^2$. *Then there exists a sequence* $\{h_n : n \in \mathbf{N}\} \subset \mathcal{L}_T^2$ *such that* $h_n(\cdot,\cdot)$ *is bounded for each* $n = 1, 2, \ldots$ *and*

$$\int_0^T E\left(|(f(t,\omega) - h_n(t,\omega)|\right)^2 dt \to 0 \text{ as } n \to \infty.$$

Proof. Put

$$h_n(t,\cdot) = \begin{cases} -n & \text{if } f(t,\omega) < -n; \\ f(t,\omega) & \text{if } |f(t,\omega)| \le n; \\ n & \text{if } f(t,\omega) > n. \end{cases}$$

Clearly, for each $t \in [0,T]$, $h_n(t,\omega) - f(t,\omega) \to 0$, a.a.[P] ω as $n \to \infty$, therefore,

$$E\left(|(f(t,\omega) - h_n(t,\omega)|\right)^2 \to 0$$

as $n \to \infty$. Moreover,

$$\int_0^T E\left(|(f(t,\omega) - h_n(t,\omega)|\right)^2 dt \le 4 \int_0^T E\left(|f(t,\cdot)|^2\right) dt < \infty.$$

The conclusion follows by the Dominated Convergence Theorem applied to the functions $E(|h_n(t,\cdot) - f(t,\cdot)|^2)$ in $L^1([0,T], \mathcal{B}_{[0,T]}, \lambda)$.

To prove Theorem 5 one has to show that for every $f \in \mathcal{L}_T^2$ there exists a sequence of step functions, $\{\phi_n : n \in \mathbf{N}\}$ in $\mathcal{S}_T^2$ such that for any given $\epsilon > 0$

$$\|f - \phi_n\|_{2,T} < \epsilon$$

for all $n \geq \bar{n}_\epsilon$. This follows from Lemmas 1, 2, and 3 and the triangle inequality for the Euclidean norm. More specifically, using the three lemmas above, one can find p_ϵ, m_ϵ, and n_ϵ such that for any $p > p_\epsilon$, $m > m_\epsilon$, $n > n_\epsilon$ we have $\|f - h_p\|_{2,T} < \epsilon/3$, $\|h_p - g_m\|_{2,T} < \epsilon/3$, and $\|\phi_n - g_m\|_{2,T} < \epsilon/3$, respectively. Using the triangle inequality we find

$$\|f - \phi_n\|_{2,T} \leq \|f - h_p\|_{2,T} + \|h_p - g_m\|_{2,T} + \|\phi_n - g_m\|_{2,T} \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

which is what we were required to prove.

For f a step function in $\mathcal{S}_T^2$ corresponding to a partition $0 = t_1 < t_2 < \ldots < t_{n+1} = T$ and random variables $f_1(\omega), f_2(\omega), \ldots, f_n(\omega)$, one can define the Itô stochastic integral of f over the interval $[0,T]$ by

$$I(f)(\omega) = \sum_{j=1}^{n} f_j(\omega) \cdot [W_{t_j+1}(\omega) - W_{t_j}(\omega)] \;\; w.p.1. \tag{1.4}$$

Theorem 6. *For any $f, g \in \mathcal{S}_T^2$ and $\alpha, \beta \in \mathbf{R}$ the Itô stochastic integral* (1.4) *satisfies*

(a) $I(f)$ is $\mathcal{F}_T$-measurable;

(b) $E[I(f)] = 0$;

*(c) $E[I(f)^2] = \int_0^T E[f(t,\omega)^2]\, dt$ (**Itô isometry**);*

(d) $I(\alpha f + \beta g) = \alpha I(f) + \beta I(g)$, w.p.1.

Proof. As $f_j(\omega)$ is $\mathcal{F}_{t_j}$-measurable, $W_{t_j+1}(\omega) - W_{t_j}(\omega)$ is $\mathcal{F}_{t_j+1}$-measurable, and $\mathcal{F}_{t_j} \subset \mathcal{F}_{t_j+1}$ for $j = 1, 2, \ldots, n$ it follows that their product $f_j(\omega) \cdot [W_{t_j+1}(\omega) - W_{t_j}(\omega)]$ is $\mathcal{F}_{t_j+1}$-measurable for $j = 1, 2, \ldots, n$. Hence $I(f)$ is $\mathcal{F}_T$-measurable.

By the Cauchy-Schwarz inequality and property (c) in Definition 4 it is easily checked that each product is integrable over Ω. Hence $I(f)$ is integrable and

$$\begin{aligned} E[I(f)] &= \sum_{j=1}^{n} E[f_j(\omega) \cdot [W_{t_j+1}(\omega) - W_{t_j}(\omega)]] \\ &= \sum_{j=1}^{n} E[E[f_j(\omega) \cdot [W_{t_j+1}(\omega) - W_{t_j}(\omega) \mid \mathcal{F}_{t_j}]]]. \end{aligned}$$

Using the fact that $W_{t_j+1}(\omega) - W_{t_j}(\omega)$ is independent of $\mathcal{F}_{t_j}$ and $f_j(\cdot)$ is $\mathcal{F}_{t_j}$-measurable, we find

$$= \sum_{j=1}^{n} E[f_j(\omega)] \cdot E[W_{t_j+1}(\omega) - W_{t_j}(\omega) \mid \mathcal{F}_{t_j}] = 0$$

since $E[W_{t_j+1}(\omega) - W_{t_j}(\omega) \mid \mathcal{F}_{t_j}] = 0$.
Now, $f_j(\cdot)$ and $f_j(\cdot) \cdot f_i(\cdot) \cdot [W_{t_i+1} - W_{t_i}]$ are also $\mathcal{F}_{t_j}$-measurable for any $i < j$. Thus,

$$E[I(f)^2] = \sum_{j=1}^{n} E\Big((f_j(\omega))^2 \cdot [W_{t_i+1}(\omega) - W_{t_i}(\omega)]^2 \Big)$$

$$+2 \sum_{j=1}^{n} \sum_{i=j+1}^{n} E\Big(f_j(\omega) \cdot f_i(\omega) \cdot [W_{t_j+1}(\omega) - W_{t_j}(\omega)] \cdot [W_{t_i+1}(\omega) - W_{t_i}(\omega)] \Big)$$

$$= \sum_{j=1}^{n} E[(f_j(\omega))^2 \cdot E\Big([W_{t_j+1}(\omega) - W_{t_j}(\omega)]^2 \Big)$$

$$+2 \sum_{j=1}^{n} \sum_{i=j+1}^{n} E\Big(f_j(\omega) \cdot f_i(\omega) \cdot [W_{t_j+1}(\omega) - W_{t_j}(\omega)] \Big) \cdot E\Big(W_{t_i+1}(\omega) - W_{t_i}(\omega) \mid \mathcal{F}_{t_j}] \Big)$$

where the law of total expectation was used. Then, since $E[(W_{t_j+1}(\omega) - W_{t_j}(\omega))^2] = t_{j+1} - t_j$ and $E[W_{t_i+1}(\omega) - W_{t_i}(\omega) \mid \mathcal{F}_{t_j}] = 0$,

$$= \sum_{j=1}^{n} E[(f_j(\omega))^2] \cdot (t_{j+1} - t_j) + 0 = \int_0^T E[(f(t,\omega))^2] dt$$

where the last equality is the consequence of the definition of the Lebesgue (or Riemann) integral for the nonrandom step function $E[(f(t,\omega))^2]$.
Finally, one should note that for $f, g \in \mathcal{S}_T^2$ and $\alpha, \beta \in \mathbf{R}$ it is also $\alpha f + \beta g \in \mathcal{S}_T^2$ with the combined step points of f and g, so by algebraic rearrangement we obtain

$$I(\alpha f + \beta g) = \alpha I(f) + \beta I(g) \;\; w.p.1.$$

The next goal is to define and establish the properties of Itô stochastic integral, $I(f)$, of functions $f \in \mathcal{L}_T^2$. The definition we adopt is the following

$$I(f)(\omega) = \lim_n I(\phi_n)(\omega) \tag{1.5}$$

where $\{\phi_n : n \in \mathbf{N}\}$ is a sequence of functions in $\mathcal{S}_T^2$, i.e. a sequence of simple functions in $\mathcal{L}_T^2$. The definition of $I(\phi_n)(\omega)$ for $\phi_n \in \mathcal{S}_T^2$ was given in (1.4). The first task is to show that $I(f)$ in (1.5) is a well defined mathematical object. To this purpose, one should note that Theorem 5 provides us with a sequence of functions $\phi_n \in \mathcal{S}_T^2$ for which

$$\int_0^T E\Big(|\phi_n(t,\omega) - f(t,\omega)|^2 \Big) dt \to 0 \text{ as } n \to \infty.$$

The Itô integrals $I(\phi_n)$ are well-defined by (1.4) and since $I(\phi_n) - I(\phi_{n+m}) = I(\phi_n - \phi_{n+m})$, they satisfy

$$E\Big(|I(\phi_n) - I(\phi_{n+m})|^2\Big)dt = E\Big(|I(\phi_n - \phi_{n+m})|^2\Big)dt$$
$$= \int_0^T E\Big(|\phi_n(t,\omega) - \phi_{n+m}(t,\omega)|^2\Big)dt$$
$$= 2\int_0^T E\Big(|\phi_n(t,\omega) - f(t,\omega) + f(t,\omega) - \phi_{n+m}(t,\omega)|^2\Big)dt$$

and, using the inequality $(a+b)^2 \le 2(a^2+b^2)$,

$$\le 2\int_0^T E\Big(|\phi_n(t,\omega) - f(t,\omega)|^2\Big)dt + 2\int_0^T E\Big(|\phi_{n+m}(t,\omega) - f(t,\omega)|^2\Big)dt.$$

As this means that $I(\phi_n)$ is a Cauchy sequence in the Banach space $L^2(\Omega,\mathcal{F},P)$, there exists a unique, w.p.1, random variable, I, in $L^2(\Omega,\mathcal{F},P)$ such that $E[|I(\phi_n) - I|^2] \to 0$ as $n \to \infty$. The random variable I is $\mathcal{F}_T$-measurable since it is the limit of $\mathcal{F}_T$-measurable random variables. Moreover, we obtain the same limit I for any choice of the sequence of step functions $\{\phi_n : n \in \mathrm{N}\}$ converging to $f \in \mathcal{L}_T^2$. In fact, let $\{\psi_n : n \in \mathrm{N}\}$ be another sequence of step functions converging to f and suppose that $I(\psi_n)$ converges to $\tilde{I}$. Then

$$E[|I - \tilde{I}|^2] \le 2E[|I - I(\phi_n)|^2] + 2E[|\tilde{I} - I(\psi_n)|^2]$$

where the second term is estimated using the same steps as above with ψ_n replacing ϕ_n. Taking limits as $n \to \infty$ we obtain $E[|I - \tilde{I}|^2] = 0$, and hence $I = \tilde{I}$ w.p.1.

We can, therefore, **define** the Itô stochastic integral, $I(f)$, of a function $f \in \mathcal{L}_T^2$ to be **the common mean-square limit** of sequences of sums (1.4) for any sequence of step functions in $\mathcal{S}_T^2$ converging to f in the norm $\|\cdot\|_{2,T}$. In addition,

common mean-square limit

Theorem 7. *The Itô stochastic integral $I(f)$ defined by* (1.5) *satisfies properties (a), (b), (c), and (d) of Theorem 6 for functions $f \in \mathcal{L}_T^2$.*

Proof. $\mathcal{F}_T$-measurability follows easily from the fact that $I(f)$ is the limit of $\mathcal{F}_T$-measurable r.v.'s, $I(\phi_n)$.

By Theorem 5 there exists a sequence of step functions $\{\phi_n : n \in \mathrm{N}\} \in \mathcal{S}_T^2$ such that

$$E[\int_0^T |\phi_n(t,\omega) - f(t,\omega)|^2\,dt] \to 0 \text{ as } n \to \infty.$$

This implies that

$$E[\int_0^T \phi_n(t,\omega)^2 dt] \to E[\int_0^T f(t,\omega)^2 dt].$$

In addition, by Theorem 6.b

$$E[\int_0^T \phi_n(t,\omega)dW_t(\omega)] = 0$$

and, by Theorem 6.c (Itô isometry),

$$E|\int_0^T \phi_n(t,\omega)dW_t(\omega)|^2 = E[\int_0^T \phi_n(t,\omega)^2 dt$$

the interchanging of the operators E and $\int$ being possible because of assumption (b) in Definition 4 and Fubini's Theorem.
Combining these facts together, it follows that

$$\begin{aligned}
E|\int_0^T \phi_n(t,\omega)dW_t(\omega) - \int_0^T \phi_m(t,\omega)dW_t(\omega)|^2 \\
&= E|\int_0^T [\phi_n(t,\omega) - \phi_m(t,\omega)]dW_t(\omega)|^2 \\
&= E\int_0^T |\phi_n(t,\omega) - \phi_m(t,\omega)|^2 dt \to 0 \text{ as } m,n \to \infty.
\end{aligned}$$

The last statement depends on the inequality

$$\begin{aligned}
&E\int_0^T |\phi_n(t,\omega) - \phi_m(t,\omega)|^2 dt \\
&\leq 2\Big(E[\int_0^T |\phi_n(t,\omega) - f(t,\omega)|^2 dt] + E[\int_0^T |\phi_m(t,\omega) - f(t,\omega)|^2 dt]\Big).
\end{aligned}$$

The result just established simply tells us that

$$\begin{aligned}
&\int_0^T \phi_n(t,\omega)dW_t(\omega) \xrightarrow{L^2} \int_0^T f(t,\omega)dW_t(\omega) \\
&\qquad \Rightarrow \int_0^T \phi_n(t,\omega)dW_t(\omega) \xrightarrow{P} \int_0^T f(t,\omega)dW_t(\omega).
\end{aligned}$$

Convergence in probability allows us to write

$$E\int_0^T f(t,\omega)dW_t(\omega) = \lim_n E\int_0^T \phi_n(t,\omega)dW_t(\omega) = 0$$

which proves part (b). From the convergence in L^2, on the other hand, we have

$$\begin{aligned}
E\Big[|\int_0^T f(t,\omega)dW_t(\omega)|^2\Big] &= \lim_n E\Big[|\int_0^T \phi_n(t,\omega)dW_t(\omega)|^2\Big] \\
&= \lim_n E\Big[\int_0^T \phi_n(t,\omega)^2\, dt\Big] \\
&= E\int_0^T f(t,\omega)^2 dt
\end{aligned}$$

thus proving (c). Part (d) can be proved easily starting from the definition of stochastic integral as the mean-square limit of sequences of the sums (1.4) in the norm $\|\cdot\|_{2,T}$.

Problem C. Prove that for any $f, g \in \mathcal{L}_T^2$

$$E[I(f)I(g)] = \int_0^T E[f(t,\omega)\cdot g(t,\omega)]dt.$$

Solution. Let $h(t,\omega) = f(t,\omega) + g(t,\omega)$. Since $f, g \in \mathcal{L}_T^2$ it is easy to verify that $h \in \mathcal{L}_T^2$ as well. This implies that we can use Theorem 7.c. and write

$$\begin{aligned}
\int_0^T E[h(t,\omega)^2]dt &= E\Big(\int_0^T h(t,\omega)dW_t(\omega)\Big)^2 \\
&= E\bigg[\Big(\int_0^T f(t,\omega)dW_t(\omega)\Big)^2 \\
&+ E\int_0^T f(t,\omega)dW_t(\omega)\cdot\int_0^T g(t,\omega)dW_t(\omega) + \Big(\int_0^T g(t,\omega)dW_t(\omega)\Big)^2\bigg] \\
&= \int_0^t E[f(t,\omega)^2]dt + \\
& 2E\int_0^T f(t,\omega)dW_t(\omega)\cdot\int_0^T g(t,\omega)dW_t(\omega) + \int_0^t E[g(t,\omega)^2]dt.
\end{aligned}$$

On the other hand, it is also

$$\begin{aligned}
\int_0^T E[h(t,\omega)^2]dt = \int_0^t E[f(t,\omega)^2]dt \\
+ 2\int_0^T E[f(t,\omega)\cdot g(t,\omega)]dt + \int_0^t E[g(t,\omega)^2]dt.
\end{aligned}$$

Comparing the two expressions for $\int_0^T E[h(t,\omega)^2]dt$ gives the desired result.

We show now that a stochastic integral is **mean-square continuous** and that it has a **separable and jointly $\mathcal{B}_{[0,T]}\times\mathcal{F}$-measurable version** which will be used from now on. **mean-square continuity** **separability**

For a variable subinterval $[t_0,t]\subseteq[0,T]$ we can form a stochastic process $Z = \{Z_t : t_0 \le t \le T\}$ defined by

$$Z_t(\omega) = \int_{t_0}^t f(s,\omega)dW_s(\omega) \tag{1.6}$$

w.p.1., for $t_0 \le t \le T$. Replacing 0 by t_0 and T by t in Theorem 7., one finds that Z_t is $\mathcal{F}_t$-measurable with $E[Z_t] = 0$ and

$$E[Z_t^2] = \int_{t_0}^t E[f(s,\omega)^2]ds \tag{1.7}$$

From Theorem 7.d. we have that for $0 \leq t_0 < t_1 < t_2 \leq T$, w.p.1.

$$\int_{t_0}^{t_2} f(s,\omega)dW_s(\omega) = \int_{t_0}^{t_1} f(s,\omega)dW_s(\omega) + \int_{t_1}^{t_2} f(s,\omega)dW_s(\omega). \tag{1.8}$$

From (1.8) and (1.6) we can write for any $0 \leq t' < t \leq T$

$$E[|Z_t - Z_{t'}|^2] = \int_{t'}^{t} E[f(s,\omega)^2]ds,$$

from which it follows that Z_t is mean-square continuous. Since Z_t is mean-square continuous it is also stochastically continuous (i.e. $P[|Z_t - Z_s| > \epsilon] \to 0$ as $s \to t$, for any $\epsilon > 0$). Then, we can use the following

separable process

Definition 8. *A stochastic process $X = \{X_t : t \in T\}$ defined on a complete probability space $(\Omega, \mathcal{F}, P)$ is called a* **separable process** *if there exists a countably dense subset $S = \{s_1, s_2, \ldots\}$ of T called a* **separant set**, *such that for any open interval I_0 and any closed interval I_c the subsets*

$$\mathcal{A} = \cup_{t \in T \cap I_0}\{\omega \in \Omega : X_t(\omega) \in I_c\} \text{ and } \mathcal{B} = \cup_{s_j \in S \cap I_0}\{\omega \in \Omega : X_{s_j}(\omega) \in I_c\}$$

differ by a subset of a null set.

Example. Let $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), \lambda)$ and on this probability space define the stochastic process $Y = \{Y_t(\omega) = I_{\{\omega\}}(t) : t \in T\}$. Then, Y is not a separable process. In fact, any countable set has zero Lebesgue measure, λ, but this is not the case for $I_0 = (\alpha, \beta)$, $0 < \alpha < \beta < 1$.

Theorem 9. *Let $\{X_t, t \in T\}$ be stochastically continuous. Then it has a separable measurable version.*

Proof. See e.g. Borkar (1995); Thm. 6.2.3.

This means that Z_t has a separable and jointly $\mathcal{B}_{[0,T]} \times \mathcal{F}$-measurable version. This makes it possible to state and prove the next important result.

Theorem 10. *A separable, jointly measurable version of Z_t defined by*

$$Z_t = \int_{t_0}^{t} f(s,\omega)dW_s(\omega)$$

for $t \in [t_0, T]$ has, almost surely, continuous sample paths.

Proof. The first step towards proving the Theorem is
Lemma A. *For $t_0 \leq s \leq t \leq T$ we have*

$$E[Z_t - Z_s \mid \mathcal{F}_s] = 0 \;\; w.p.1.$$

Proof. Let $Z_t^{(n)}(\omega) = \int_{t_0}^{t_1} \phi^{(n)}(s,\omega)dW_s(\omega)$, $n = 1, 2, \ldots$ and $\phi^{(n)} \in \mathcal{S}_T^2$. Then, we have w.p.1.

$$Z_t^{(n)}(\omega) - Z_s^{(n)}(\omega) = \int_s^t \phi^{(n)}(u,\omega)dW_u(\omega).$$

Let introduce the following partition $s = t_1^{(n)} < t_2^{(n)} < \ldots < t_{n+1}^{(n)} = t$ of $[s,t]$ and let $\phi_j^{(n)}(t_j^{(n)}, \omega) = \phi_j^{(n)}(\omega)$. Then, we have that $\phi_j^{(n)}(\omega)$ is $\mathcal{F}_{t_j^{(n)}}$-measurable, whereas $E[W_{t_{j+1}^{(n)}} - W_{t_j^{(n)}} \mid \mathcal{F}_{t_j^{(n)}}] = 0$ w.p.1, for $j = 1, 2, \ldots n$. Hence

$$E[\phi_j^{(n)}[W_{t_{j+1}^{(n)}} - W_{t_j^{(n)}}] \mid \mathcal{F}_{t_j^{(n)}}] = \phi_j^{(n)} E[W_{t_{j+1}^{(n)}} - W_{t_j^{(n)}} \mid \mathcal{F}_{t_j^{(n)}}] = 0,$$

w.p.1, for $j = 1, 2, \ldots, n$. Now,

$$E[Z_t^{(n)}(\omega) - Z_s^{(n)}(\omega) \mid \mathcal{F}_s] = E[\sum_{j=1}^{n} \phi_j^{(n)}(\omega)[W_{t_{j+1}^{(n)}}(\omega) - W_{t_j^{(n)}}(\omega) \mid \mathcal{F}_s]$$

and, since $\mathcal{F}_s \subseteq \mathcal{F}_{t_1^{(n)}} \subseteq \mathcal{F}_{t_2^{(n)}} \subseteq \ldots$,

$$= E[E[\sum_{j=1}^{n} \phi_j^{(n)}(\omega)[W_{t_{j+1}^{(n)}}(\omega) - W_{t_j^{(n)}}(\omega)] \mid \mathcal{F}_{t_j^{(n)}}] \mid \mathcal{F}_s]$$

$\phi_j^{(n)}(\omega)$ is $\mathcal{F}_{t_j^{(n)}}$-adapted

$$= E[\sum_{j=1}^{n} \phi_j^{(n)}(\omega) E[W_{t_{j+1}^{(n)}}(\omega) - W_{t_j^{(n)}}(\omega) \mid \mathcal{F}_{t_j^{(n)}}] \mid \mathcal{F}_s] = 0.$$

Since we know that by our definition of stochastic integral

$$Z_t^{(n)}(\omega) - Z_s^{(n)}(\omega) \xrightarrow{L^2} Z_t(\omega) - Z_s(\omega),$$

we can state that[5]

$$E[Z_t(\omega) - Z_s(\omega) \mid \mathcal{F}_s]^2 = \lim_n E[Z_t^{(n)}(\omega) - Z_s^{(n)}(\omega) \mid \mathcal{F}_s]^2 = 0$$

from which the statement of Lemma A follows.

We have then shown that $\{Z_t : t \in [0,T]\}$ **is a separable martingale with finite second moment.** To prove its continuity, we can use the fact that $\{Z_t : t \in [0,T]\}$ is a separable martingale and use the martingale maximal inequality

$$P[\sup_{t_0 \le s \le t} |Z_t| \ge a] \le \frac{1}{a^2} \int_{t_0}^{t} E[f(s,\omega)^2] ds \tag{1.9}$$

for any $a > 0$, and the Doob Inequality (with $p = 2$):

$$E\left[\sup_{t_0 \le s \le t} |Z_t|^2\right] \le 4E[Z_t^2] = 4 \int_{t_0}^{t} E[f(s,\omega)^2] ds. \tag{1.10}$$

[5]We are using the fact that if $X_n \xrightarrow{L^2} X$, then $E[X_n \mid \mathcal{G}] \xrightarrow{L^2} E[X \mid \mathcal{G}]$ where $\mathcal{G}$ is a relevant σ-algebra for these random variables.

The difference of two martingales with respect to the same increasing family of σ-algebras is itself a martingale. Hence for $f \in \mathcal{L}_T^2$ and $\{\phi^{(n)}\}$ a sequence of step functions converging to f in $\mathcal{L}_T^2$, the difference $Z_t - Z_t^{(n)}$ is a martingale, where both Z_t and $Z_t^{(n)}$ were defined above for f and $\phi^{(n)}$, respectively. Then, by (1.9) and choosing the step functions $\phi^{(n)}$ so that

$$\int_{t_0}^{T} E\Big(|f(s,\omega) - \phi^{(n)}(s,\omega)|^2\Big)ds \le \frac{1}{n^4},$$

we have

$$P\Big(\sup_{t_0\le s\le T} |Z_t - Z_t^{(n)}| \ge \frac{1}{n}\Big) \le n^2 \int_{t_0}^{T} E\Big(|f(s,\omega) - \phi^{(n)}(s,\omega)|^2\Big)ds \le \frac{1}{n^2}$$

for $n = 1, 2, 3, \ldots$ Therefore,

$$\sum_{n=1}^{\infty} P\Big(\sup_{t_0\le s\le T} |Z_t - Z_t^{(n)}| \ge \frac{1}{n}\Big)$$

is convergent and by the Borel-Cantelli Lemma

$$N = \cap_{n\ge 1} \cup_{k\ge n} \Big\{\omega \in \Omega : \sup_{t_0\le s\le T} |Z_t(\omega) - Z_t^{(n)}(\omega)| \ge \frac{1}{n}\Big\}$$

has measure 0. In other words, we have established that

$$\sup_{t_0\le s\le T} |Z_t(\omega) - Z_t^{(n)}(\omega)| \ge \frac{1}{n}$$

can occur for at most a finite number of n. This proves that

$$\lim_{n} \sup_{t_0\le s\le T} |Z_t(\omega) - Z_t^{(n)}(\omega)| = 0.$$

Since $Z_t(\omega)$ is the uniform limit in $t \in [0, T]$ of continuous functions $Z_t^{(n)}(\omega)$ it is also continuous.

4. Computing Stochastic Integrals

In the previous section we have provided a definition of the mathematical objects that we refer to as Itô stochastic integrals, i.e. the L^2-limit

$$\int_0^T \phi^{(n)}(t,w)dW_t(\omega) \xrightarrow{L^2} I_T(f) = \int_0^T f(t,\omega)dW_t(\omega),$$

or

$$\lim_{n} E\Big|I_T(f) - \int_0^T \phi^{(n)}(t,w)dW_t(\omega)\Big|^2 = 0$$

where $\phi^{(n)}(t,w)$ is a suitable step functions such that

$$\lim_n \int_0^T E[f(t,\omega) - \phi^{(n)}(t,\omega)]^2 dt = 0.$$

In principle this provides the practitioner with two ways to compute stochastic integrals $\int_0^T f(t,\omega)dW_t$:

i. guess, $I_T(f)$, and verify that

$$\lim_n E\left|I_T(f) - \sum_{j=0}^{n-1} f(t_j,\omega)\Delta W_{j+1}(\omega)\right|^2 = 0,$$

where $\Delta W_{j+1}(\omega) = W_{t_{j+1}}(\omega) - W_{t_j}(\omega)$.

ii. find a sequence of bounded step functions for which

$$\lim_n \int_0^T E[f(t,\omega) - \phi^{(n)}(t,\omega)]^2 dt = 0$$

and then, compute the integral as

$$\int_0^T f(t,\omega)dW_t = \lim_n \sum_{j=0}^{n} \phi^{(n)}(t_j,\omega)\cdot[W_{t_{j+1}}(\omega) - W_{t_j}(\omega)].$$

Problem D. Compute $\int_0^T t\,dW_t$.

Solution. For any partition $0 = t_0 < t_1 < t_2 < \ldots < t_n = T$, such that $\Delta t_{j+1} \equiv t_{j+1} - t_j \to 0$ as $n \to \infty$, let $\phi^{(n)}(t,\omega) = t_j \cdot I_{[t_j,t_{j+1})}(t)$. In this case where the step functions $\phi^{(n)}(t,\omega)$ are non-stochastic it suffices to verify the condition

$$\lim_n \int_0^T [\phi^{(n)}(t,\omega) - t]^2 dt = 0.$$

Now,

$$\int_0^T [\phi^{(n)}(t,\omega) - t]^2 dt = \sum_{j=1}^{n} \int_{t_j}^{t_j+1} (t_j - t)^2 \cdot dt = \frac{1}{3}\sum_{j=1}^{n} (\Delta t_{j+1})^3 \to 0.$$

as $n \to \infty$. The sequence of step functions above clearly satisfies the requirement and one is able to compute the stochastic integral above as

$$\begin{aligned}
\int_0^T t\,dW_t &= \lim_n \int_0^T \phi^{(n)}(t)dW_t \\
&= \lim_n \sum_{j=1}^{n} \phi_j^{(n)} \Delta W_{j+1} \\
&= \lim_n \left(T\cdot W_T - \sum_{j=0}^{n-1} W_{t_j}\cdot\Delta t_{j+1}\right) = T\cdot W_T - \int_0^T W_t\,dt.
\end{aligned}$$

Thus we have established that

$$\int_0^T t\,dW_t = T \cdot W_T - \int_0^T W_t\,dt.$$

Problem E. Show that $\int_0^T W_s dW_s = \frac{1}{2}W_T^2 - \frac{T}{2}$.

Solution. To check this statement we can write the approximating sum, $\sum_j W_j \Delta W_{j+1}$, and check whether

$$\lim_n E\left|\frac{1}{2}W_T^2 - \frac{T}{2} - \sum_{j=0}^{n-1} W_j \Delta W_{j+1}\right|^2 = 0$$

or not. To this purpose, let

$$A_n \equiv E\left|\frac{1}{2}W_T^2 - \sum_{j=0}^{n-1} W_j \Delta W_{j+1}\right|^2$$

and

$$B_n \equiv E\left[\left(\frac{1}{2}W_T^2 - \sum_{j=0}^{n-1} W_j \Delta W_{j+1}\right) - \frac{T}{2}\right]^2.$$

Now,

$$\begin{aligned}
A_n &= E\left|\frac{1}{2}W_T^2 - \sum_{j=0}^{n-1} W_j \Delta W_{j+1}\right|^2 \\
&= E\left[\frac{W_T^4}{4} + \left(\sum_{j=0}^{n-1} W_j \Delta W_{j+1}\right)^2 - W_T^2 \cdot \sum_{j=0}^{n-1} W_j \Delta W_{j+1}\right] \\
&= \frac{3T^2}{4} + \sum_{j=0}^{n-1} t_j \Delta t_{j+1} - 2\sum_{j=0}^{n-1} t_j \Delta t_{j+1} \\
&\to \frac{3 \cdot T^2}{4} - \int_0^T t dt = \frac{T^2}{4}.
\end{aligned}$$

where we have used the facts

(a) $\left(\sum_{j=0}^{n-1} W_j \Delta W_{j+1}\right)^2 = \sum_{j=0}^{n-1} W_j^2 \Delta W_{j+1}^2 + 2\sum_{i<j} W_i \cdot W_j \cdot \Delta W_{i+1} \cdot \Delta W_{j+1}$,

(b)

$$\sum_{i<j} W_i \cdot W_j \cdot \Delta W_{i+1} \cdot \Delta W_{j+1}$$
$$= \sum_{i<j} W_i^2 \cdot [W_j - W_{i+1} + \Delta W_{i+1} + W_i] \cdot \Delta W_{i+1} \cdot \Delta W_{j+1}$$
$$= \sum_{i<j} W_i \cdot [W_j - W_{i+1}] \cdot \Delta W_{i+1} \cdot \Delta W_{j+1}$$
$$+ W_i \cdot (\Delta W_{i+1})^2 \cdot \Delta W_{j+1}$$
$$+ W_i^2 \cdot \Delta W_{i+1} \cdot \Delta W_{j+1},$$

(c) $W_T = (W_T - W_{j+1}) + \Delta W_{j+1} + W_j,\ j = 0, 1, \ldots, n-1.$

(d) the stochastic independence of W_i, ΔW_{i+1}, $(W_j - W_{i+1})$ and ΔW_{j+1}.

In particular, from (b) and (d) it follows that $E[2\sum_{i<j} W_i \cdot W_j \cdot \Delta W_{i+1} \cdot \Delta W_{j+1}] = 0$, from (c) we find that

$$W_T^2 \cdot \sum_{j=0}^{n-1} W_j \Delta W_{j+1} = \sum_{j=0}^{n-1} \left[(W_T - W_{j+1}) + \Delta W_{j+1} + W_j\right]^2 \cdot W_j \cdot \Delta W_{j+1}$$

and, hence, using (d), it follows that

$$E[W_T^2 \cdot \sum_{j=0}^{n-1} W_j \Delta W_{j+1}] = \sum_{j=0}^{n-1} E[W_j^2] \cdot E[\Delta W_{j+1}^2] = \sum_{j=0}^{n-1} t_j \cdot \Delta t_{j+1}$$

which proves our statement. Therefore,

$$B_n = A_n - T(A_n)^{1/2} + \frac{T^2}{4} \to 0$$

as $n \to \infty$ and the claim is verified.

Problem F. Compute $\int_0^T W_t^2 dW_t$.

Solution. Let $0 = t_0 < t_1 < t_2 < \ldots < t_n = T$ be any partition of the interval $[0, T]$ such that $\Delta t_j \to 0$ as $n \to \infty$ for any $j = 1, 2, \ldots, n$. We identify a step function $\phi^{(n)} \in \mathcal{L}_T^2$ such that $\int_0^T E|\phi^{(n)}(t,\omega) - W_t^2(\omega)|^2\,dt \to 0$ as $n \to \infty$. A reasonable choice is

$$\phi^{(n)}(t,\omega) = W_{t_j}^2(\omega) \cdot I_{[t_j, t_{j+1})}(t).$$

In fact,

$$\int_0^T E|\phi^{(n)}(t,\omega) - W_t^2(\omega)|^2\,dt$$
$$= \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} E[(W_{t_j}^2 - W_t)^2]\,dt = \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} E[W_{t_j}^4 + W_t^4 - 2W_{t_j}^2 W_t^2]\,dt$$

and, since W_t can also be written as $(W_t - W_{t_j}) + W_{t_j}$, writing W_t in the form $(W_t - W_j) + W_j$, the properties of Brownian motion and a few algebraic manipulations make it possible to rewrite the last term above as

$$= \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} (3t^2 - t_j^2 - 2t_j t)dt = \sum_{j=0}^{n-1} (\Delta t_{j+1})^2 (t_{j+1} + t_j)$$
$$\leq 2T \cdot \sum_{j=0}^{n-1} (\Delta t_{j+1})^2 \to 0$$

as $n \to \infty$. This proves that we can compute the stochastic integral at hand as

$$\lim_n \sum_{j=0}^{n-1} W_j^2 \cdot \Delta W_{j+1}$$

given that the limit is independent of the particular sequence $\phi^{(n)}(\cdot, \cdot)$ used. Now,

$$W_T^3 = \sum_j \Delta W_{j+1}^3 = \sum_j (\Delta W_{j+1})^3 + 3W_{j+1}^2 W_j - 3W_j^2 W_{j+1}$$
$$= \sum_j (\Delta W_{j+1})^3 + 3W_j W_{j+1} \Delta W_{j+1}$$
$$= \sum_j (\Delta W_{j+1})^3 + 3 \cdot \sum_j W_j (\Delta W_{j+1})^2 + \sum_j \sum_j W_j^2 \Delta W_{j+1}$$

or, rearranging the terms,

$$\sum_j \sum_j W_j^2 \Delta W_{j+1} = \frac{W_T^3}{3} - \sum_j W_j (\Delta W_{j+1})^2 - \frac{1}{3} \sum_j (\Delta W_{j+1})^3.$$

It is not difficult to show that when $n \to \infty$,

$$\sum_j (\Delta W_{j+1})^3 \xrightarrow{L^2} 0$$

and

$$\sum_j W_j (\Delta W_{j+1})^2 \xrightarrow{L^2} \sum_j W_j \Delta t_{j+1}.$$

Therefore, we have established that

$$\sum_j^{n-1} W_j^2 \Delta W_{j+1}$$
$$= \frac{W_T^3}{3} - \sum_j^{n-1} W_j \cdot [(\Delta W_{j+1})^2 - \Delta t_{j+1}] + \sum_j^{n-1} W_j \cdot \Delta t_{j+1} + \sum_j^{n-1} (\Delta W_{j+1})^3$$
$$\xrightarrow{L^2} \frac{W_T^3}{3} - \int_0^T W_t \, dt.$$

The exercises above show that the basic definition of stochastic integrals is not very convenient in evaluating a given integrals. In some sense, the basic definition of Itô integrals is not very useful when one tries to evaluate a given stochastic integral. This is similar to the situation for ordinary Riemann integrals, where we do not use the basic definition but rather the fundamental theorem of calculus plus the chain rule in the explicit calculations. In the context of stochastic integrals, however, there is no differentiation theory, only integration theory. Nevertheless it turns out that it is possible to establish an Itô integral version of the chain rule, the Itô **formula**. Before devoting our attention to this formula, we will show how to introduce a more general version of Itô stochastic integral.

5. A More General Definition of Itô Integral

The definition of stochastic integral in section 2 was obtained requiring that the random functions, $f(t,\omega)$, belongs to the class $\mathcal{L}^2_T$. Most commonly encountered random functions fulfill this requirement. Nevertheless, it is possible to extend the definition of stochastic integral to a larger class of integrands. An example of a random function that does not belong to $\mathcal{L}^2_T$ is

$$f(t,\omega) \equiv \exp\{c \cdot W_t^2(\omega)\}$$

where as usual $W_t(\omega)$ is the Wiener process and $T > (2c)^{-1}$. In fact, it is easy to show that

$$\int_0^T E|f(t,\omega)|^2\,dt = \infty.$$

Despite this difficulty, it is still possible to provide a meaningful definition of stochastic integral even for functions not in $\mathcal{L}^2_T$. To this purpose let $\mathcal{L}^2_\omega[\alpha,\beta]$ be the class of functions such that

(a) $f(t,\omega)$ is a separable process;

(b) $f(t,\omega)$ is a $\mathcal{B}_{[\alpha,\beta]} \times \mathcal{F}$-measurable process;

(c) $f(t,\omega)$ is $\mathcal{F}_t$-adapted for any $t \in [\alpha,\beta]$;

(d) $P[\int_\alpha^\beta |f(t,\omega)|^2\,dt < \infty] = 1$.

Clearly, $\mathcal{L}^2[\alpha,\beta] \subset \mathcal{L}^2_\omega[\alpha,\beta]$. The starting point towards the definition of a new stochastic integral is the following

Theorem 11. *Let $f \in \mathcal{L}^2_\omega[\alpha,\beta]$. Then,*

i. there exists a sequence of continuous functions $g_n \in \mathcal{L}^2_\omega[\alpha,\beta]$ such that

$$\lim_n \int_\alpha^\beta |f(t,\omega) - g_n(t,\omega)|^2 dt = 0 \;\; a.s.;$$

ii. there exists a sequence of step functions $\psi_n \in \mathcal{L}^2_\omega[\alpha,\beta]$ such that

$$\lim_n \int_\alpha^\beta |f(t,\omega) - \psi_n(t,\omega)|^2 dt = 0 \ a.s.;$$

Proof. The proof provided here is very similar to that in Friedman (1975). Let $r(t)$ be defined by

$$r(t) = \begin{cases} \exp\{1/(t^2-1)\} & \text{if } |t| \le 1; \\ 0 & \text{if } |t| > 1. \end{cases}$$

Define $f(t) = 0$ if $t < \alpha$ and let

$$(I_\epsilon f)(t) = \frac{1}{2r(0)\epsilon} \int_{\alpha-1}^\beta r\left(\frac{t-s-\epsilon}{\epsilon}\right) f(s)ds, \ \ \epsilon < 1/2.$$

Then $(I_\epsilon f)(t)$ is a continuous function in t since

$$\lim_{h\to 0} |I_\epsilon f)(t+h) - I_\epsilon f)(t)|$$
$$\le \lim_{h\to 0} \frac{1}{2r(0)\epsilon} \int_{\alpha-1}^\beta \left| r\left(\frac{t+h-s-\epsilon}{\epsilon}\right) - r\left(\frac{t-s-\epsilon}{\epsilon}\right)\right| f(s)ds$$

and, since by the DCT for nonrandom functions we can interchange the operations of limit and integration,

$$= \frac{1}{2r(0)\epsilon} \int_{\alpha-1}^\beta \lim_{h\to 0} \left| r\left(\frac{t+h-s-\epsilon}{\epsilon}\right) - r\left(\frac{t-s-\epsilon}{\epsilon}\right)\right| f(s)ds = 0$$

using the continuity of $r(\cdot)$. In addition, from the definition of $r(\cdot)$ we can write

$$(I_\epsilon f)(t) = \frac{1}{2r(0)\epsilon} \int_{t-2\epsilon}^t r\left(\frac{t-s-\epsilon}{\epsilon}\right) f(s)ds = \frac{1}{2r(0)\epsilon} \int_{-\epsilon}^\epsilon r\left(\frac{z}{\epsilon}\right) f(t-z-\epsilon)dz$$

which is the consequence of requiring $-1 \le (t-s-\epsilon)/\epsilon \le 1$. Also, by Scharwz's Inequality, we have that

$$\int_\alpha^\beta [(I_\epsilon f)(t)]^2 \le \int_\alpha^\beta f^2(t)dt. \tag{1.11}$$

This follows from the following argument

$$\int_\alpha^\beta (I_\epsilon f)^2 dt = \int_\alpha^\beta \left[\int_{-\epsilon}^\epsilon \frac{1}{2r(0)\epsilon} r(z/\epsilon) f(s-z-\epsilon)dz\right]^2 ds$$
$$\le \int_\alpha^\beta \left[\int_{-\epsilon}^\epsilon \frac{1}{2r(0)\epsilon} r(z/\epsilon)dz \cdot \int_{-\epsilon}^\epsilon \frac{1}{2r(0)\epsilon} r(z/\epsilon) f^2(s-z-\epsilon)dz\right] ds.$$

Since $r(z/\epsilon) \leq r(0)$, we can transform the previous inequality as

$$\int_\alpha^\beta (I_\epsilon f)^2 dt \leq \int_\alpha^\beta \left[\int_{-\epsilon}^{\epsilon} \frac{1}{2r(0)\epsilon} r(z/\epsilon) dz \int_{-\epsilon}^{\epsilon} f^2(s-z-\epsilon) dz \right] ds.$$

Then, interchanging the order of integration,

$$\begin{aligned}\int_\alpha^\beta (I_\epsilon f)^2 dt &\leq \int_{-\epsilon}^{\epsilon} \frac{1}{2r(0)\epsilon} r(z/\epsilon) \left[\int_\alpha^\beta f^2(s-z-\epsilon) ds \right] dz \\ &\leq \int_{-\epsilon}^{\epsilon} \frac{1}{2r(0)\epsilon} r(z/\epsilon) \left[\int_{\alpha-z-\epsilon}^{\beta-z-\epsilon} f^2(t) dt \right] dz\end{aligned}$$

now, since $\epsilon < 1/2$, we have that $\beta - z - \epsilon < \beta$, and $\alpha - z - \epsilon > \alpha - 1$ so that we the inequality can be rewritten as

$$\int_\alpha^\beta (I_\epsilon f)^2 dt \leq \int_{-\epsilon}^{\epsilon} \frac{1}{2r(0)\epsilon} r(z/\epsilon) \left[\int_{\alpha-1}^{\beta} f^2(t) dt \right] dz$$

and, as it is easily seen that $\int_{-\epsilon}^{\epsilon} \frac{1}{2r(0)\epsilon} r(z/\epsilon) dz \leq 1$, we have found

$$\int_\alpha^\beta (I_\epsilon f)^2 dt \leq \int_{\alpha-1}^{\beta} f^2(t) dt = \int_\alpha^\beta f^2(t) dt.$$

For fixed ω for which $\int_\alpha^\beta f^2(t,\omega) dt < \infty$, let $\{u_n : n \in \mathrm{N}\}$ be a sequence of nonrandom continuous functions such that $u_n(t,\omega) = 0$ if $t < \alpha$ and

$$\int_\alpha^\beta |u_n(t) - f(t,\omega)|^2 \, dt \to 0 \tag{1.12}$$

if $n \to \infty$. Using the generalized first mean value theorem and the continuity of $u_n(\cdot)$, it is possible to show that

$$(I_\epsilon f)(t) \to u_n(t) \text{ uniformly } \quad t \in [\alpha, \beta]$$

as $\epsilon \to 0$. Now, it is easily checked that $(I_\epsilon f)(t)(I_\epsilon u_n)(t) = (I_\epsilon(f - u_n))(t)$, and this fact can be exploited to write

$$\begin{aligned}\int_\alpha^\beta |(I_\epsilon f)(t,\omega) - f(t,\omega))|^2 dt \leq& \int_\alpha^\beta |(I_\epsilon(f(t,\omega) - u_n(t))|^2 dt \\ &+ \int_\alpha^\beta |(I_\epsilon u_n)(t) - u_n(t)|^2 dt + \int_\alpha^\beta |u_n(t) - f(t,\omega)|^2 dt\end{aligned}$$

and using (1.11) with f replaced by $f - u_n$ together with the fact that, by construction, $(I_\epsilon f)(t) \to u_n(t)$ uniformly in $[\alpha, \beta]$ when $\epsilon \to 0$,

$$\limsup_{\epsilon \to 0} \int_\alpha^\beta |(I_\epsilon f)(t,\omega) - f(t,\omega))|^2 dt \leq 2 \int_\alpha^\beta |u_n(t) - f(t,\omega)|^2 dt.$$

Since by assumption $f \in \mathcal{L}^2_\omega[\alpha,\beta]$, this result holds for almost all ω and therefore, taking $n \to \infty$ and using (1.12), we have established that

$$\limsup_{\epsilon \to 0} \int_\alpha^\beta |(I_\epsilon f)(t,\omega) - f(t,\omega))|^2 dt = 0 \ a.s.$$

Thus, the first statement of the Theorem holds with $g_n = (I_{1/n} f)$.
To prove the second statement, let

$$h_{n,m}(t) = \begin{cases} g_n(k/m) & \text{if } \alpha + k/m \le t < \alpha + (k+1)/m; \\ 0 & \text{otherwise} \end{cases}$$

with $0 \le l < m(\beta - \alpha)$. Then,

$$\lim_m \int_\alpha^\beta |h_{n,m}(t) - g_n(t)| dt = 0 \ a.s. \tag{1.13}$$

Now, for any $\delta > 0$, there exists n_0 such that $\forall n \ge n_0$

$$P\left[\int_\alpha^\beta |h_{n,m}(t) - g_n(t)|^2 dt > \frac{\delta}{2}\right] < \frac{\delta}{2}.$$

From (1.13) and letting $n = n_0$ one finds that there exists also an m_0 for which

$$P\left[\int_\alpha^\beta |g_{n_0}(t) - h_{n_0,m}(t)|^2 dt > \frac{\delta}{2}\right] < \frac{\delta}{2}$$

$\forall m \ge m_0$. Hence,

$$P\left[\int_\alpha^\beta |f(t) - h_{n_0,m_0}(t)|^2 dt > \delta\right] < \delta.$$

Taking $\delta = 1/k$ and denoting the corresponding h_{n_0,m_0} by ψ_k, it follows that $\psi_k \in \mathcal{L}^2_\omega[\alpha,\beta]$ and

$$\int_\alpha^\beta |f(t) - \psi_k(t)|^2 dt \xrightarrow{P} 0.$$

But then there is a subsequence $\{f_n : n \in \mathbf{N}\}$ of $\{\psi_k; k \in \mathbf{n}\}$ satisfying (ii).

Definition of stochastic integral in $\mathcal{L}^2_\omega[\alpha,\beta]$. Let $f(t)$ be a step function in $\mathcal{L}^2_\omega[\alpha,\beta]$, say $f(t) = f_i$ if $t_i \le t < t_{i+1}$, $0 \le i \le r-1$ where $\alpha = t_0 < t_1 < \ldots < t_r = \beta$. The random variable

$$\sum_{k=0}^{r-1} f(t_k,\omega) \cdot [W_{t_k+1}(\omega) - W_{t_k}(\omega)]$$

is denoted by

$$\int_\alpha^\beta f(t,\omega) dW_t(\omega)$$

and is called the stochastic integral of $f(\cdot,\cdot)$ with respect to the Brownian motion $W(\cdot,\cdot)$; it is also called the Itô integral. It is easy to prove from the definition above

Theorem 12. *Let f_1, f_2 be two step functions in $\mathcal{L}^2_\omega[\alpha,\beta]$ and let λ_1, λ_2 be real numbers. Then $\lambda_1 f_1 + \lambda_2 f_2$ is in $\mathcal{L}^2_\omega[\alpha,\beta]$ and*

$$\int_\alpha^\beta [\lambda_1 f_1(t,\omega) + \lambda_2 f_2(t,\omega)]dW_t(\omega)$$
$$= \lambda_1 \int_\alpha^\beta f_1(t,\omega)dW_t(\omega) + \lambda_2 \int_\alpha^\beta f_2(t,\omega)dW_t(\omega).$$

Theorem 13. *For any step function $f \in \mathcal{L}^2_\omega[\alpha,\beta]$ and for any $\epsilon > 0, N > 0$,*

$$P\Big[\Big|\int_\alpha^\beta f(t,\omega)dW_t(\omega)\Big| > \epsilon\Big] \le P\Big[\int_\alpha^\beta f(t,\omega)dt > N\Big] + \frac{N}{\epsilon^2}. \tag{1.14}$$

Proof. Let

$$\phi_N(t,\omega) = \begin{cases} f(t,\omega) & \text{if } t_k \le t < t_{k+1} \text{ and } \sum_{j=0}^k f^2(t_j,\omega)\Delta_{t_{j+1}} \le N; \\ 0 & \text{if } t_k \le t < t_{k+1} \text{ and } \sum_{j=0}^k f^2(t_j,\omega)\Delta_{t_{j+1}} > N, \end{cases}$$

where $f(t) = f(t_j)$ if $t_j \le t < t_{j+1}$; $t_0 = \alpha < t_1 < \ldots < t_r = \beta$. Then, $\phi_N \in \mathcal{L}^2_\omega[\alpha,\beta]$ and, if ν is the largest integer such that $\nu \le r-1$ and

$$\sum_{j=0}^{\nu} f^2(t_j,\omega)\Delta_{t_{j+1}} \le N,$$

then

$$\int_\alpha^\beta \phi_N^2(t,\omega)\,dt = \sum_{j=0}^{\nu} f^2(t_j,\omega)\Delta_{t_{j+1}}.$$

Hence,

$$E\Big[\int_\alpha^\beta \phi_N^2(t,\omega)\,dt\Big] \le N.$$

Now, since $f(t,\omega) = \phi_N(t,\omega)$ for all $t \in [\alpha,\beta)$ if $\int_\alpha^\beta f^2(t,\omega)dt < N$, we have that

$$
\begin{aligned}
P&\left[\left\{\omega : \left|\int_\alpha^\beta f(t,\omega)\,dW_t(\omega)\right| > \epsilon\right\}\right] \\
&= P\left[\left\{\omega : \left|\int_\alpha^\beta f(t,\omega)\,dW_t(\omega)\right| > \epsilon\right\} \cap \left\{\omega : \int_\alpha^\beta f^2(t,\omega)\,dt \le N\right\}\right] \\
&+ P\left[\left\{\omega : \left|\int_\alpha^\beta f(t,\omega)\,dW_t(\omega)\right| > \epsilon\right\} \cap \left\{\omega : \int_\alpha^\beta f^2(t,\omega)\,dt > N\right\}\right] \\
&\le P\left[\left\{\omega : \left|\int_\alpha^\beta \phi_N(t,\omega)\,dW_t(\omega)\right| > \epsilon\right\}\right] + P\left[\left\{\omega : \int_\alpha^\beta f^2(t,\omega)\,dt > N\right\}\right].
\end{aligned}
$$

Clearly, by construction, $\phi_N \in \mathcal{L}^2[\alpha,\beta]$ and, hence, the Itô isometry allows one to write

$$E[\int_\alpha^\beta \phi_N(t,\omega)\,dW_t(\omega)] = E[\int_\alpha^\beta \phi_N(t,\omega)\,dt].$$

Now, using Chebyshev's inequality, we find that

$$P\left[\left|\int_\alpha^\beta f(t,\omega)\,dW_t(\omega)\right| > \epsilon\right] \le \frac{1}{\epsilon^2} E\left|\int_\alpha^\beta \phi_N(t,\omega)\,dW_t(\omega)\right|^2 \le \frac{N}{\epsilon^2}$$

which proves the Theorem.

We are finally in the condition to define the stochastic integral for any function $f \in \mathcal{L}^2_\omega[\alpha,\beta]$. In fact, by Theorem 11 there is a sequence of step functions $f_n \in \mathcal{L}^2_\omega[\alpha,\beta]$ such that for $n \to \infty$

$$\int_\alpha^\beta |f_n(t,\omega) - f(t,\omega)|^2\,dt \xrightarrow{P} 0. \tag{1.15}$$

Hence,

$$\lim_{m,n\to\infty} \int_\alpha^\beta |f_n(t,\omega) - f_m(t,\omega)|^2\,dt \xrightarrow{P} 0.$$

By Theorem 13, for any $\epsilon > 0$, and $\rho > 0$, (let $\rho = N/\epsilon^2$), we have

$$
\begin{aligned}
P&\left[\left|\int_\alpha^\beta f_n(t,\omega)dW_t(\omega) - \int_\alpha^\beta f_m(t,\omega)dW_t(\omega)\right| > \epsilon\right] \\
&\le \rho + P\left[\left|\int_\alpha^\beta f_n(t,\omega) - f_m(t,\omega)\right|^2 dt > \epsilon^2\rho\right]
\end{aligned}
$$

and, therefore, it follows that the sequence

$$\left\{\int_\alpha^\beta f_n(t,\omega)dW_t(\omega) : n \in \mathbf{N}\right\}$$

is **convergent in probability**. The limit, denoted by $\int_\alpha^\beta f(t,\omega)dW_t(\omega)$ is the stochastic integral (Itô integral) of $f(t,\omega)$ with respect to the Brownian motion $W_t(\omega)$.

Before we are done, we need to show that the definition is independent of the particular sequence of step functions $\{f_n : n \in \mathbf{N}\}$. To this purpose, let $\{g_n : n \in \mathbf{N}\}$ be another sequence of step functions in $\mathcal{L}^2_\omega[\alpha,\beta]$ converging to f in the sense that

$$\int_\alpha^\beta |g_n(t,\omega) - g(t,\omega)|^2 dt \xrightarrow{P} 0.$$

Define now the new sequence $\{h_n : n \in \mathbf{N}\}$ as

$$h_n(t,\omega) = \begin{cases} f_n(t,\omega) & \text{if n is even;} \\ g_n(t,\omega) & \text{if n is odd.} \end{cases}$$

Then, by what it was proved above, the sequence

$$\left\{\int_\alpha^\beta h_n(t,\omega)\, dW_t(\omega) : n \in \mathbf{N}\right\}$$

is convergent in probability. On the other hand, it contains two convergent subsequences, $\{\int_\alpha^\beta h_{2n}(t,\omega)\, dW_t(\omega) : n \in \mathbf{N}\}$ and $\{\int_\alpha^\beta h_{2n+1}(t,\omega)\, dW_t(\omega) : n \in \mathbf{N}\}$, respectively. Then, the limits in probability of $\int_\alpha^\beta f_n dW$ and $\int_\alpha^\beta g_n dW$ are equal a.s.

The final task is to extend the definition to any function in $\mathcal{L}^2_\omega[\alpha,\beta]$. The following result is easily checked

Theorem 14. *Let f_1, f_2 be funtions from $\mathcal{L}^2_\omega[\alpha,\beta]$ and let λ_1, λ_2 be real numbers. Then $\lambda_1 f_1 + \lambda_2 f_2$ is in $\mathcal{L}^2_\omega[\alpha,\beta]$ and*

$$\int_\alpha^\beta [\lambda_1 f_1(t,\omega)+\lambda_2 f_2(t,\omega)]\, dW_t(\omega) = \lambda_1 \int_\alpha^\beta f_1(t,\omega)dW_t(\omega)+\lambda_2 \int_\alpha^\beta f_2(t,\omega)dW_t(\omega).$$

The next result extends Theorem 13 to any function in $\mathcal{L}^2_\omega[\alpha,\beta]$

Theorem 15. *If f is any function from $\mathcal{L}^2_\omega[\alpha,\beta]$, then for any $\epsilon > 0, N > 0$,*

$$P\left[|\int_\alpha^\beta f(t,\omega)dW_t(\omega)|> \epsilon\right] \leq P\left[\int_\alpha^\beta f(t,\omega)dt > N\right] + \frac{N}{\epsilon^2}. \tag{1.16}$$

Proof. By Theorem 11 there exists a sequence of step functions $f_n \in \mathcal{L}^2_\omega[\alpha,\beta]$ such that

$$\int_\alpha^\beta |f_n(t,\omega) - f(t,\omega)|^2 dt \xrightarrow{P} 0. \tag{1.17}$$

By definition of the stochastic integral,

$$\int_\alpha^\beta f_n(t,\omega)dW_t(\omega) \xrightarrow{P} \int_\alpha^\beta f(t,\omega)dW_t(\omega) \tag{1.18}$$

as $n \to \infty$. Applying Theorem 12 to f_n we get

$$P\left[|\int_\alpha^\beta f_n(t,\omega)dW_t(\omega)| > \epsilon'\right] \leq P\left[\int_\alpha^\beta f(t,\omega)dt > N'\right] + \frac{N'}{(\epsilon')^2}.$$

Taking $n \to \infty$ and using (1.17) and (1.18), we find that for any $\epsilon > \epsilon'$ and $N < N'$, it is

$$P\left[|\int_\alpha^\beta f(t,\omega)dW_t(\omega)| > \epsilon\right] \leq P\left[\int_\alpha^\beta f(t,\omega)dt > N\right] + \frac{N}{\epsilon^2}.$$

Now, letting $\epsilon' \uparrow \epsilon$, and $N' \downarrow N$ gives (1.16).

The results proved so far only show that it is possible to define the object we call stochastic integral as the limit in probability of some sequence of stochastic integrals for step integrands. What we need is a practical method to compute stochastic integrals. This is done in the following

Theorem 16. *If $f \in \mathcal{L}^2_\omega[\alpha,\beta]$ and f is continuous, then, for any sequence of partitions $\alpha = t_{n,0} < t_{n,1} < \ldots < t_{n,m_n} = \beta$ of $[\alpha,\beta]$ such that*

$$\lim_n \max_{1 \leq j < m_n} \Delta t_{n,j+1} = 0,$$

$$\sum_{k=0}^{m_n-1} f(t_{n,k},\omega)\Delta W_{t_{n,k+1}}(\omega) \xrightarrow{P} \int_\alpha^\beta f(t,\omega)dW_t(\omega) \tag{1.19}$$

as $n \to \infty$.

Proof. We need first the following

Lemma. *Let f, f_n be in $\mathcal{L}^2_\omega[\alpha,\beta]$ and suppose that*

$$\int_\alpha^\beta |f_n(t,\omega) - f(t,\omega)|^2 dt \xrightarrow{P} 0 \ \text{ as } n \to \infty. \tag{1.20}$$

Then,

$$\int_\alpha^\beta f_n(t,\omega)dW_t(\omega) \xrightarrow{P} \int_\alpha^\beta f(t,\omega)dW_t(\omega) \ \text{ as } n \to \infty. \tag{1.21}$$

Proof. By Theorem 15, for any $\epsilon > 0, \rho > 0$, we have that

$$P\left[\left|\int_\alpha^\beta (f_n(t,\omega) - f(t,\omega)dW_t(\omega)\right| > \epsilon\right] \leq P\left[\int_\alpha^\beta |f_n(t,\omega) - f(t,\omega)|^2 dt > \epsilon^2\rho\right] + \rho.$$

Taking $n \to \infty$ and using (1.20), completes the proof of the lemma.

Now, let the step function g_n be defined by

$$g_n(t,\omega) = f(t_{n,k},\omega)$$

if $t_{n,k} \leq t < t_{n,k+1}$, $0 \leq k \leq m_n - 1$. Then, for a.a. ω, $g_n(t,\omega) \to f(t,\omega)$ uniformly in $t \in [\alpha,\beta)$ as $n \to \infty$. Hence,

$$\int_\alpha^\beta |g_n(t,\omega) - f(t,\omega)|^2 dt \to 0 \text{ a.s.}$$

By the lemma we just proved, this implies that

$$\int_\alpha^\beta g_n(t,\omega)dW_t(\omega) \xrightarrow{P} \int_\alpha^\beta f(t,\omega)dW_t(\omega).$$

The assertion of the theorem follows from

$$\int_\alpha^\beta g_n(t,\omega)dW_t(\omega) = \sum_{k=0}^{m_n-1} f(t_{n,k},\omega)\Delta W_{t_{n,k+1}}(\omega).$$

Problem G. Let $f,g \in \mathcal{L}^2_\omega[\alpha,\beta]$ and assume that $f(t,\omega) = g(t,\omega)$ for all $t \in [\alpha,\beta]$, $\omega \in \Omega_0$. Then, prove that

$$\int_\alpha^\beta f(t,\omega)dW_t(\omega) = \int_\alpha^\beta g(t,\omega)dW_t(\omega) \ \textit{for a.a. } \omega \in \Omega_0.$$

Solution. Let $\{\psi_k : k \in \mathbf{N}\}$ be a sequence of step functions in $\mathcal{L}^2_\omega[\alpha,\beta]$ as in the proof of Theorem 11, i.e. a sequence of step functions such that

$$\int_\alpha^\beta |f(t,\omega) - \psi_k(t,\omega)|^2 dt \xrightarrow{P} 0.$$

Similarly, let $\{\phi_k : k \in \mathbf{N}\}$ be a sequence of step functions such that

$$\int_\alpha^\beta |g(t,\omega) - \phi_k(t,\omega)|^2 dt \xrightarrow{P} 0.$$

From the construction in Theorem 10 one deduce that it is possible to choose the sequences above so that, if $\omega \in \Omega_0$, $\psi_k(t,\omega) = \phi_k(t,\omega)$ for $\alpha \leq t \leq \beta$. Hence, by the definition of the stochastic integral of a step function,

$$\int_\alpha^\beta \psi_k(t,\omega)dW_t(\omega) = \int_\alpha^\beta \phi_k(t,\omega)dW_t(\omega) \text{ if } \omega \in \Omega_0.$$

Now, the assertion follows letting $k \to \infty$.

When $f \in \mathcal{L}^2[0,T]$ we have seen (Theorem 10) that

$$Z_t(\omega) = \int_0^t f(s,\omega)dW_s(\omega)$$

is a martingale and admits a separable, jointly measurable version which has almost surely continuous sample paths. Despite the integral is the convergence in probability of mean-square intgrable martingales, when $f \in \mathcal{L}^2_\omega[0,T]$ we lose the martingale property but we can still find a continuous version of the stochastic process $Z_t(\omega)$.

Theorem 17. *If $f \in \mathcal{L}^2_\omega[0,T]$, then the integral $Z_t(\omega)$, $t \in [0,T]$ has a continuous version.*

Proof. See Friedman (1975) pp. 67-9.

We now state three results whose proofs can be found in A. Friedman, *op. cit.*.

Theorem 18. *For any $f \in \mathcal{L}^2_\omega[0,T]$ and for any α, β, γ such that $0 \le \alpha < \beta < \gamma \le T$*

$$\int_\alpha^\gamma f(s,\omega)dW_s(\omega) = \int_\alpha^\beta f(s,\omega)dW_s(\omega) + \int_\beta^\gamma f(s,\omega)dW_s(\omega) \tag{1.22}$$

Theorem 19. *Let $f \in \mathcal{L}^2_\omega[0,T]$. Then, for any $\epsilon > 0, N > 0$,*

$$P\left[\sup_{0\le t\le T}\left|\int_0^t f(s,\omega)dW_s(\omega)\right| > \epsilon\right] \le P\left[\int_0^t f^2(s,\omega)dt > N\right] + \frac{N}{\epsilon^2}. \tag{1.23}$$

Theorem 20. *Let $f_n, f \in \mathcal{L}^2_\omega[0,T]$, and assume that $\int_0^T |f_n - f|^2 dt \xrightarrow{P} 0$ if $n \to \infty$. Then, if $n \to \infty$,*

$$\sup_{0\le t\le T}\left|\int_0^t f_n(s,\omega)dW_s(\omega) - \int_0^t f(s,\omega)dW_s(\omega)\right| \xrightarrow{P} 0 \tag{1.24}$$

The final remark is about the requirement $P[\int_\alpha^\beta |f(t,\omega)|^2\,dt < \infty] = 1$. In particular, we will show that this requirement cannot be removed. In fact, assume that $[\alpha, \beta] = [0,1]$ and, furthermore, we assume that f is non-stochastic and such that

$$P[\int_0^t f(s)^2\,dt < \infty,\ t < 1] = 1 \text{ and } P[\int_0^1 f(s)^2\,dt = \infty] = 1.$$

Set $\tau(0) = 0$, and for $0 < t < 1$, let $\tau(t) = \int_0^t f^2(u)du$. Now, $\tau(t)$ is a non-decreasing function and as such it admits a left-continuous inverse function, $\tau^{-1}(t) = \min\{s : \tau(s) = t\}$, with $\tau^{-1}(0) = 0$. For any $t \in [0,1)$, define

$$Y_{\tau(t)} = Z_{\tau^{-1}(t)} = \int_0^{\tau^{-1}(t)} f(s)dW_s(\omega).$$

Clearly, we have that $f \in \mathcal{L}^2[0,1]$ and it is easy to show that Y_τ is a Wiener process. In first place $\tilde{\mathcal{F}}_\tau = \mathcal{F}_{\tau^{-1}(t)}$, Y_τ is $\tilde{\mathcal{F}}_\tau$-measurable with

$$Y_0 = 0, \quad E[Y_{\tau(t)} - Y_{\tau(s)} \mid \tilde{\mathcal{F}}_\tau(s)] = 0$$

and

$$E[|Y_{\tau(t)} - Y_{\tau(s)}|^2 | \tilde{\mathcal{F}}_\tau(s)] = \int_s^t f^2(u)du = \tau(t) - \tau(s)$$

w.p.1. Hence by a theorem of Doob, the process $\{Y_\tau, \tau > 0\}$ is a Wiener process with respect to the family of $\sigma-$algebras $\{\tilde{\mathcal{F}}_\tau, \tau \geq 0\}$, at least for $0 \leq \tau < \tau(1)$. Using the time substitution (intrinsic time clock for Z), one finds that the stochastic integral

$$Y_{\tau(t)} = \int_0^{\tau^{-1}(t)} f(s)\, dW_s(\omega)$$

is a Brownian motion. In addition, by assumption, $\lim_{t\to 1} \tau(t) = \infty$, w.p.1. But, for a Brownian motion we know that

$$\limsup_{\tau(t)\to\infty} Y_{\tau(t)} = -\liminf_{\tau(t)\to\infty} Y_{\tau(t)} = \infty$$

and, hence, we have found that

$$P[\limsup_{t\uparrow 1} \int_0^t f(s)dW_s(\omega) = -\liminf_{t\uparrow 1} \int_0^t f(s)dW_s(\omega) = \infty] = 1.$$

Therefore, the condition $P[\int_0^1 f(t,\omega)^2\, dt < \infty] = 1$ is indispensable for the existence of $\int_0^1 f(t,\omega)dW_t(\omega)$.

6. Itô Formula

Definition 21. *Let X_t, $0 \leq t \leq T$ be a process such that for any $0 \leq t_1 < t_2 \leq T$*

$$X_{t_2} - X_{t_1} = \int_{t_1}^{t_2} a(t)dt + \int_{t_1}^{t_2} b(t)dW_t$$

where $a \in \mathcal{L}^1_\omega[0,T]$ and $b \in \mathcal{L}^2_\omega[0,T]$. Then we say that X_t has stochastic differential dX, on $[0,T]$, given by

$$dX_t = a(t)dt + b(t)dW_t.$$

In particular, one should observe that X_t is a nonanticipative function, a continuous process and belongs to $\mathcal{L}^\infty_\omega[0,T]$.

Example 1. If $0 \leq t_1 < t_2$ and $\max \Delta t_{n,j+1} \to 0$ as $n \to \infty$ for any sequence of partitions of $[t_1, t_2] : t_1 = t_{n,1}, t_{n,2}, \cdots, t_{n,n} = t_2$ then, by Theorem 16, we have that

$$\begin{aligned}\int_{t_1}^{t_2} W_t dW_t &= \lim_n \sum_{k=1}^{n-1} W_{t_{n,k}} \Delta W_{t_{n,k+1}} \\ &= \frac{1}{2} \lim_n \sum_{k=1}^{n-1} [W_{t_{n,k+1}}^2 - W_{t_{n,k}}^2] - (\Delta W_{t_{n,k+1}})^2 \\ &= \frac{1}{2} \cdot [W_{t_2}^2 - W_1^2 - \frac{1}{2} \cdot \lim_n \sum_{k=1}^{n-1} (\Delta W_{t_{n,k+1}})^2\end{aligned}$$

where $\lim_n$ is taken as the limit in probability. For a Wiener process, the last limit in probability is equal to $t_2 - t_1$. Therefore,

$$\int_{t_1}^{t_2} W_t dW_t = \frac{1}{2} \cdot [W_{t_2}^2 - W_{t_1}^2] - \frac{1}{2} \cdot (t_2 - t_1), \tag{1.25}$$

or

$$d(W_t^2) = dt + 2W_t\, dW_t. \tag{1.26}$$

Example 2. By Theorem 16,

$$\int_{t_1}^{t_2} t dW_t = \lim_n \sum_{k=1}^{n-1} t_{n,k} \Delta W_{t_{n,k+1}} \text{ in probability.}$$

In addition, it is clearly

$$\int_{t_1}^{t_2} W_t dt = \lim_n \sum_{k=1}^{n-1} W_{t_{n,k+1}} \cdot (t_{n,k+1} - t_{n,k})$$

for all ω for which $W_t(\omega)$ is continuous. The sum of the right-hand sides is equal to

$$\lim_n \sum_{k=1}^{n-1} [t_{n,k+1} W_{t_{n,k+1}} - t_{n,k} W_{t_{n,k}}] = t_2 W_{t_2} - t_1 W_{t_1}$$

or

$$d(tW_t) = W_t\, dt + t\, dW_t. \tag{1.27}$$

Theorem 22. *If* $dX_i(t,\omega) = a_i(t,\omega) + b_i(t,\omega)\, dW_t(\omega)$, $i = 1, 2$, *then*

$$d(X_1(t,\omega) \cdot X_2(t,\omega)) = X_1(t,\omega) dX_2(t,\omega) + X_2(t,\omega) dX_1(t,\omega) + b_1(t,\omega) b_2(t,\omega) dt. \tag{1.28}$$

The integrated form of 1.28 states that, for any $0 \le t_1 < t_2 \le T$,

$$
\begin{aligned}
X_1(t,\omega)\cdot X_2(t,\omega) = &\int_{t_1}^{t_2} X_1(t,\omega)a_2(t,\omega)dt + \int_{t_1}^{t_2} X_1(t,\omega)b_2(t,\omega)dt + \\
&+ \int_{t_1}^{t_2} X_2(t,\omega)a_1(t,\omega)dt + \int_{t_1}^{t_2} X_2(t,\omega)b_1(t,\omega)dt \qquad (1.29)\\
&+ \int_{t_1}^{t_2} b_1(t,\omega)b_2(t,\omega)dt.
\end{aligned}
$$

Proof. In the following, to ease the notation, the argument for ω is suppressed. Suppose first that a_i, b_i, $i = 1,2$ are constants in the interval $[t_1, t_2)$. Then, (1.29) follows from examples 1 and 2. Next, if a_i, b_i, $i = 1,2$ are step functions in $[t_1, t_2)$, constants on successive intervals $I_1, I_2, \dots, I_l$, then (1.29) holds with t_1 and t_2 replaced by the end points of each interval I_i, $i = 1, 2, \dots, l$. Taking the sum gives (1.29).
Consider now the general case. Approximate a_i, b_i, $i = 1,2$ by nonanticipative step functions $a_{i,n}$, $b_{i,n}$ in such a way that

$$\int_0^T |a_{i,n}(t) - a_i(t)|dt \to 0 \text{ a.s.},$$

$$\int_0^T |b_{i,n}(t) - b_i(t)|^2 dt \to 0 \text{ a.s.}$$

Let

$$X_{i,n}(t) = X_i(0) + \int_0^t a_{i,n}(s)ds + \int_0^t b_{i,n}(s)dW_s.$$

By Theorem 22, the second condition implies that

$$\sup_{0\le t\le T} |\int_0^t b_{i,n}(s)dW_s - \int_0^t b_i(s)dW_s| \xrightarrow{P} 0 \text{ if } n \to \infty.$$

Hence, from

$$|X_{i,n}(t) - X_i(t)| = |\int_0^t a_{i,n}(s) - a_i(s)ds + \int_0^t b_{i,n}(s) - b_i(s)dW_s|$$

one finds that

$$
\begin{aligned}
&\sup_{0\le t\le T} |X_{i,n}(t) - X_i(t)| \le \\
&\quad \sup_{0\le t\le T} \int_0^t |a_{i,n}(s) - a_i(s)|ds + \sup_{0\le t\le T} |\int_0^t b_{i,n}(s)dW_s - \int_0^t b_i(s)dW_s| \xrightarrow{P} 0
\end{aligned}
$$

as $n \to \infty$ because of the assumption about the $a_{i,n}$ and the implication proved above. Then, this implies that there is a subsequence $\{X_{i,n_k}(t)\}$ such that

$$X_{i,n_k}(t) \to X_i(t) \text{ uniformly in } [0,T], \text{ a.s.} \qquad (1.30)$$

Using (1.30) and the Lemma used in the proof of Theorem 16 one can easily check that it is

$$\int_{t_1}^{t_2} X_{i,n_k}(t)b_{j,n}(t)dW_t \xrightarrow{P} \int_{t_1}^{t_2} X_i(t)b_j(t)dW_t \text{ as } k \to \infty.$$

Similarly, one proves also

$$\int_{t_1}^{t_2} X_{i,n}(t)a_{j,n}(t)dW_t \to \int_{t_1}^{t_2} X_i(t)a_j(t)dW_t$$

and

$$\int_{t_1}^{t_2} b_{1,n}(t)b_{2,n}(t)dt \to \int_{t_1}^{t_2} b_1(t)b_2(t)dt$$

a.s. as $n \to \infty$. Writing (1.29) for $a_{i,n}, b_{i,n}, X_{i,n}$ and letting $n \to \infty$, proves the assertion (1.29). Then, as t_1 and t_2 are arbitrary, the proof of the theorem is complete.

Lemma. *Let $U : [0,T] \times \mathbf{R} \to \mathbf{R}$ have continuous partial derivatives $\frac{\partial U}{\partial t}$, $\frac{\partial U}{\partial x}$ and $\frac{\partial^2 U}{\partial x^2}$. Then, for any t, $t+\Delta t \in [0,T]$ and x, $x+\Delta x \in \mathbf{R}$ there exist constants $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$ such that*

$$\begin{aligned} &U(t+\Delta t, x+\Delta x) - U(t,x) = \\ &\quad \frac{\partial U}{\partial t}(t+\alpha\Delta t, x) + \frac{\partial U}{\partial x}(t,x)\cdot\Delta x + \frac{1}{2}\cdot\frac{\partial^2 U}{\partial x^2}(t, x+\beta\Delta x)(\Delta x)^2. \end{aligned} \tag{1.31}$$

Proof. Use Taylor and Mean Value Theorems of classical calculus.

Writing a_t and b_t for $a(t,\omega)$ and $b(t,\omega)$, respectively, we are now ready for

Theorem 23. (Itô's Formula) *Let $Y_t = U(t, X_t)$ for $0 \leq s \leq t \leq T$ where U is as in the Lemma above and X_t satisfies*

$$X_t - X_s = \int_s^t a_t\, du + \int_s^t b_t\, dW_u,$$

with $\sqrt{|a|}, b \in \mathcal{L}^2_\omega[0,T]$. Then

$$\begin{aligned} &Y_t - Y_s = \\ &\quad \int_s^t \left[\frac{\partial U}{\partial t}(u, X_u) + a_u\cdot\frac{\partial U}{\partial x}(u, X_u) + \frac{1}{2}\cdot b_u^2\cdot\frac{\partial^2 U}{\partial x^2}(u, X_u)\right] du \\ &\quad + \int_s^t b_u\cdot\frac{\partial U}{\partial x}(u, X_u)\, dW_u \end{aligned} \tag{1.32}$$

w.p.1, for any $0 \leq S \leq t \leq T$.

Proof.

Step 1. Suppose that a and b do not depend on t, so they are $\mathcal{F}_0$-measurable random variables. We choose a sample-path continuous version of X_t (this is possible because of Theorem 17) and fix a subinterval $[s,t] \subseteq [0,T]$, of which we consider partitions of the form $s = t_{1,n}, t_{2,n}, \dots, t_{n+1,n} = t$ with $\Delta t_{j,n} = t_{j+1,n} - t_{j,n}$. Then,

$$Y_t - Y_s = U(t, X_t) - U(s, X_s) = \sum_{j=1}^{n} \Delta U_{j,n}$$

where

$$\Delta U_{j,n} = U(t_{j+1,n}, X_{t_{j+1,n}}) - U(t_{j,n}, X_{t_{j,n}})$$

for $j = 1, 2, \dots n$. Applying the Lemma on each subinterval $[t_{j,n}, t_{j+1,n}]$ for each $\omega \in \Omega$ we have $\alpha_{j,n}(\omega), \beta_{j,n}(\omega) \in [0,1]$ such that

$$\begin{aligned}\Delta U_{j,n} = {} & \frac{\partial U}{\partial t}\Big(t_{j,n} + \alpha_{j,n}\Delta t_{j,n}, X_{t_{j,n}}\Big)\cdot\Delta t_{j,n} \\ & + \frac{\partial U}{\partial x}\Big(t_{j,n}, X_{t_{j,n}}\Big)\Delta X_{j,n} \\ & + \frac{1}{2}\frac{\partial^2 U}{\partial x^2}\Big(t_{j,n}, X_{t_{j,n}} + \beta_{j,n}\Delta X_{j,n}\Big)\cdot(\Delta X_{j,n})^2,\end{aligned} \tag{1.33}$$

w.p.1, where $\Delta X_{j,n} = X_{t_{j+1,n}} - X_{t_{j,n}}$ for $j = 1, 2, \dots, n$. By the continuity of $\frac{\partial U}{\partial t}$, $\frac{\partial^2 U}{\partial x^2}$, and the sample-path continuity of X_t, we have for each $j = 1, 2, \dots, n$

$$\frac{\partial U}{\partial t}\Big(t_{j,n} + \alpha_{j,n}\Delta t_{j,n}, X_{t_{j,n}}\Big) - \frac{\partial U}{\partial t}\Big(t_{j,n}, X_{t_{j,n}}\Big) \to 0 \text{ w.p.1}, \tag{1.34}$$

and

$$\frac{\partial^2 U}{\partial x^2}\Big(t_{j,n}, X_{t_{j,n}} + \beta_{j,n}\Delta X_{j,n}\Big) - \frac{\partial^2 U}{\partial x^2}\Big(t_{j,n}, X_{t_{j,n}}\Big) \to 0 \text{ w.p.1}, \tag{1.35}$$

where $\max_{1\le j\le n} \Delta t_{j,n} \to 0$ as $n \to \infty$.
For a and b independent of t, the increments are of the form $\Delta X_{j,n} = a\Delta t_{j,n} + b\Delta W_{j,n}$ where $\Delta W_{j,n} = W_{t_{j+1,n}} - W_{t_{j,n}}$ for $j = 1, 2, \dots, n$. Hence

$$\begin{aligned} & \sum_{j=1}^{n}\Big[(\Delta X_{j,n})^2 - (f\Delta W_{j,n})^2\Big] \\ = {} & a^2 \sum_{j=1}^{n}\Big[(\Delta t_{j,n})^2 + 2ab\sum_{j=1}^{n} \Delta W_{j,n}\Delta t_{j,n},\end{aligned} \tag{1.36}$$

which tends to 0 in probability for $\max_{1\le j\le n} \Delta t_{j,n} \to 0$ as $n \to \infty$. Combining (1.33), (1.34), (1.35), and (1.36) and letting $\delta_n = \max_{1\le j\le n} \Delta t_{j,n}$, one finds that

$$
\begin{aligned}
Y_t - Y_s &= \lim_{\delta_n \to 0, n\to\infty} \sum_{j=1}^{n} \Delta U_{j,n} \\
&= \lim_{\delta_n \to 0, n\to\infty} \sum_{j=1}^{n} \bigg[\frac{\partial U}{\partial t}(t_{j,n}, X_{t_{j,n}}) + a\frac{\partial U}{\partial x}(t_{j,n}, X_{t_{j,n}}) \\
&+ \frac{1}{2}b^2 \frac{\partial^2 U}{\partial x^2}(t_{j,n}, X_{t_{j,n}})\bigg] \cdot \Delta t_{j,n} \\
&+ \lim_{\delta_n \to 0, n\to\infty} \sum_{j=1}^{n} b\frac{\partial U}{\partial x}(t_{j,n}, X_{t_{j,n}})\Delta W_{j,n} \\
&+ \lim_{\delta_n \to 0, n\to\infty} \sum_{j=1}^{n} \frac{1}{2}b^2 \frac{\partial^2 U}{\partial x^2}(t_{j,n}, X_{t_{j,n}}) \cdot \Big((\Delta W_{j,n})^2 - \Delta t_{j,n} \Big).
\end{aligned}
\tag{1.37}
$$

The first two limits on the right side of (1.37) are the terms on the right side of (1.32), so one has to prove that the last limit vanishes.
To this purpose, let $\Gamma_{j,n} = (\Delta W_{j,n})^2 - \Delta t_{j,n}$ and $I^N_{j,n} = I_{A^N_{j,n}}$, i.e. the indicator function of the set

$$A^N_{j,n} = \{\omega \in \Omega : |X_{t_{j,n}}| \leq N \text{ for } i = 1, 2, \dots, j\}$$

for $j = 1, 2, \dots, n$. For fixed n the random variables $\Gamma_{j,n}$ are independent and satisfy

$$E[\Gamma_{j,n}] = 0 \text{ and } E[(\Gamma_{j,n})^2] = 2(\Delta t_{j,n})^2.$$

This implies that

$$
\begin{aligned}
E&\left[\left|\sum_{j=1}^{n} \frac{\partial^2 U}{\partial x^2}(t_{j,n}, X_{t_{j,n}}) I^N_{j,n} \cdot \Gamma_{j,n}\right|\right] \\
&= \sum_{j=1}^{n} E\left[\left|\frac{\partial^2 U}{\partial x^2}(t_{j,n}, X_{t_{j,n}}) I^N_{j,n} \cdot \Gamma_{j,n}\right|\right] \\
&\leq C\sum_{j=1}^{n} 2((\Delta t_{j,n})^2 \leq 2C|t-s| \cdot \max_{1\leq j\leq n} \Delta t_{j_n} \to 0
\end{aligned}
\tag{1.38}
$$

as $n \to \infty$, where

$$C = \max_{s\leq u\leq t, |x|\leq N} \left|\frac{\partial^2 U}{\partial x^2}(u, x)\right|^2 < \infty.$$

This implies that the last limit in (1.37) vanishes if we can prove that $P(A^N_{j,n}) \to 1$ as $N \to \infty$. This follows from the fact that

$$\cup_{j=1}^{n} (A^N_{j,n})^c \subseteq B^N = \{\omega \in \Omega : \sup_{s\leq u\leq t} |X_u(\omega)| > N\},$$

and $P(B^N) \to 0$ as $N \to \infty$. In fact, using the definition of $X(t,\omega)$, it is easily checked [6] that

$$P[\sup_{s\le u\le t}\int_s^u |a|\,du > N] + P[\sup_{s\le u\le t}|\int_s^u b\,dW_u|] \ge P[\sup_{s\le u\le t}|X_u| > N],$$

and the assertion follows using an argument similar to that in Theorem 15. This completes the proof for functions a and b which do not depend on t.

Step 2. If a and b are step functions, the proof is similar as these do not vary within common partition subintervals.

Step 3. For general integrands a and b with $\sqrt{|a|}, b \in \mathcal{L}^2_\omega[0,T]$ one can (see the proof of Theorem 22.) find sequences of step functions $\{a_n : n \in \mathbf{N}\}$ and $\{b_n : n \in \mathbf{N}\}$ in $\mathcal{L}^2[0,T]$ such that the integrals

$$\int_s^t \left|a_n(u,\omega) - a(u,\omega)\right| du, \quad \int_s^t \left|b_n(u,\omega) - b(u,\omega)\right|^2 du$$

converge a.s. to zero. Thus, the sequence defined by

$$X_{n,r} = X_s + \int_s^r a_n(u)\,du + \int_s^r b_n(u)\,dW_u$$

converges in probability to X_r as $n \to \infty$ for each $s \le r \le t$. By taking subsequences, if necessary, but retaining the original index for simplicity, we can replace each of these convergences in probability by convergence with probability one; moreover this can be done uniformly on the interval $[s,t]$. As Itô's formula holds for step functions we have

$$\begin{aligned}
Y_{t,n} - Y_{s,n} &= U(t,X_{t,n}) - U(s,X_{s,n}) \\
&+ \int_s^t \left[\frac{\partial U}{\partial t}(u,X_{u,n}) + a_n(u)\frac{\partial U}{\partial x}(u,X_{u,n}) + \frac{1}{2}(b_n(u))^2\frac{\partial^2 U}{\partial x^2}(u,X_{u,n})\right] du \\
&+ \int_s^t b_n(u)\frac{\partial U}{\partial x}(u,X_{u,n})\,dW_u,
\end{aligned} \tag{1.39}$$

w.p.1, for each n. Now, $X_{u,n} \xrightarrow{P} X_u$ as $n \to \infty$. By the triangle inequality it follows for convergence in probability that

$$\begin{aligned}
&\int_s^t \left[\frac{\partial U}{\partial t}(u,X_{u,n}) + a_n(u)\frac{\partial U}{\partial x}(u,X_{u,n}) + \frac{1}{2}(b_n(u))^2\frac{\partial^2 U}{\partial x^2}(u,X_{u,n})\right] du \\
&\qquad \to \int_s^t \left[\frac{\partial U}{\partial t}(u,X_u) + a(u)\frac{\partial U}{\partial x}(u,X_u) + \frac{1}{2}(b(u))^2\frac{\partial^2 U}{\partial x^2}(u,X_u)\right] du
\end{aligned}$$

[6] Use the following arguments:

$$P[C] + P[D] \ge P[C] + P[D] - P[C \cap D] = P[C \cup D],$$

and let $C = \{\omega : \sup_{s\le u\le t}\int_s^u |a|dv > N\}$ and $D = \{\omega : \sup_{s\le u\le t}|\int_s^u b^2\,dW_v| > N\}$.

and

$$\int_s^t b_n(u)\frac{\partial U}{\partial x}(u, X_{u,n})\, dW_u \to \int_s^t b(u)\frac{\partial U}{\partial x}(u, X_u) dW_u.$$

In fact, taking subsequences if necessary, these can be considered to hold with probability one. In fact, each path of the process X_t is continuous, w.p.1, and thus bounded, w.p.1. This means that for each path all the terms appearing in (1.39) are bounded, so one can apply the Lebesgue Dominated Convergence Theorem to each continuous sample path to conclude that the first integral in (1.39) converges, w.p.1, to the first integral in (1.32). The second convergence follows from

$$\int_s^t \left| b_n(u)\frac{\partial U}{\partial x}(u, X_{u,n}) - b(u)\frac{\partial U}{\partial x}(u, X_u)\right|^2 du \xrightarrow{P} 0$$

and Theorem 20. To this purpose, one should notice that

$$\begin{aligned}
&\int_s^t \left| b_n(u)\frac{\partial U}{\partial x}(u, X_{u,n}) - b(u)\frac{\partial U}{\partial x}(u, X_u)\right|^2 du \\
&\quad \leq \int_s^t \left| (b_n(u) - b(u)) \cdot \frac{\partial U}{\partial x}(u, X_{u,n}) + b(u) \cdot \left(\frac{\partial U}{\partial x}(u, X_u) - \frac{\partial U}{\partial x}(u, X_{u,n})\right)\right|^2 du \\
&\quad \leq 2\left[\int_s^t (b_n(u) - b(u))^2 \cdot \left(\frac{\partial U}{\partial x}(u, X_{u,n})\right)^2 du \right. \\
&\qquad \left. + \int_s^t b^2(u) \cdot \left(\frac{\partial U}{\partial x}(u, X_u) - \frac{\partial U}{\partial x}(u, X_{u,n})\right)^2 du\right].
\end{aligned}$$

Now, each path of the process X_t is continuous, w.p.1. (Theorem 17), and thus bounded w.p.1. This means that the terms involving $\frac{\partial U}{\partial x}$ are bounded for each path $X_{u,n}$. The first integral is bounded by $\max_{s\leq u\leq t}|\frac{\partial U}{\partial x}|^2 \int_s^t (b_n(u) - b(u))^2 du$ which is integrable as by assumption $b, b_n \in \mathcal{L}^2_\omega[0, T]$. The second term is bounded above by $\max_{s\leq u\leq t}|\frac{\partial U}{\partial x}(u, X_u) - \frac{\partial U}{\partial x}(u, X_{u,n})|^2 \cdot \int_s^t b^2(u) du$ which is also integrable. This allows us to use the DCT and, since $X_{n,u} \xrightarrow{P} X_u$ as $n \to \infty$, every subsequence $\{X_{n,k;u}\}$ has a further subsequence $\{X_{n,k,j;u}\}$ such that $X_{n,k,j;u} \xrightarrow{a.s.} X_u$ as $j \to \infty$. Thus,

$$\int_s^t \left| b_{n,k,j}(u)\frac{\partial U}{\partial x}(u, X_{n,k,j;u}) - b(u)\frac{\partial U}{\partial x}(u, X_u)\right|^2 du \xrightarrow{a.s.} 0.$$

This establishes that the sequence $\{\int_s^t |b_n(u)\frac{\partial U}{\partial x}(u, X_{u,n}) - b(u)\frac{\partial U}{\partial x}(u, X_{u,n})\}$ has a subsequence which has a further subsequence $\{\int_s^t b_{n,k,j}(u)\frac{\partial U}{\partial x}(u, X_{n,k,j;u}) - b(u)\frac{\partial U}{\partial x}(u, X_u)\}$ which converge a.s. to zero as $j \to \infty$. This proves that

$$\int_s^t \left| b_n(u)\frac{\partial U}{\partial x}(u, X_{u,n}) - b(u)\frac{\partial U}{\partial x}(u, X_u)\right|^2 du \xrightarrow{P} 0.$$

The proof is therefore complete.

Problem H. Let $dX_t = f_t dW_t$, $f_t \in \mathcal{L}^2_\omega[0,T]$ and $Y_t = U(t,X_t)$ with $U(t,X_t) = e^{X_t}$. Prove that

$$dY_t = \frac{1}{2} f_t^2 Y_t dt + f_t Y_t dW_t.$$

Solution. With this choice for $U(t,X_t)$ one finds that

$$\frac{\partial U}{\partial t}(t,X_t) = 0; \quad \frac{\partial U}{\partial x}(t,X_t) = e^{X_t}; \quad \frac{\partial^2 U}{\partial x^2}(t,X_t) = e^{X_t}.$$

Now, the Itô's formula (with $a_t \equiv 0$ and $b_t = f_t$), gives

$$\begin{aligned} dY_t &= \frac{1}{2} f_t^2 e^{X_t} dt + f_t e^{X_t} dW_t \\ &\equiv \frac{1}{2} f_t^2 U(t,X_t) dt + f_t U(t,X_t) dW_t \\ &\equiv \frac{1}{2} f_t^2 Y_t dt + f_t Y_t dW_t. \end{aligned}$$

as we were supposed to show.

Problem I. Use Itô's formula to prove that

$$\int_0^t W_s^2 dW_s = \frac{1}{3} W_t^3 - \int_0^t W_s ds$$

where W_t is a standard Brownian motion starting at 0 w.p.1.

Solution. Let $X_t = W_t$ and $Y_t = U(t,X_t) = (1/3) \cdot W_t^3$. Now, we can always write

$$dX_t \equiv dW_t = 0\, dt + 1\, dW_t.$$

Using Itô's formula with $a_t \equiv 0$ and $b_t \equiv 1$ we find that

$$dY_t = W_t dt + W_t^2 dW_t$$

or, in integral form,

$$Y_t - Y_0 = \int_0^t W_s ds + \int_0^t W_s^2 dW_s.$$

Since $W_0 = 0$ w.p.1. we have $Y_0 = 0$, w.p.1. and rearranging the terms gives the desired result.

Problem J. Suppose $f(s,\omega) = f(s)$ only depends of s and f is continuously differentiable and of bounded variation in $[0,t]$. Prove that

$$\int_0^t f(s)\, dW_s = f(t) W_t - \int_0^t W_s\, df_s.$$

Solution. Let $X_t = W_t$ and this time set $Y_t = U(t, X_t) = f(t) \cdot W_t$. Then, because of the assumptions about the deterministic function $f(\cdot)$ we can use Itô's formula:

$$Y_t - Y_0 = \int_0^t \frac{df(s)}{ds} \cdot W_s \, ds + \int_0^t f(s) \, dW_s.$$

Assuming again that W_t is the standard Brownian motion starting at 0, we have $Y_0 = 0$ and $Y_t = f(t) \cdot W_t$. Simple algebraic manipulations then give

$$\int_0^t f(s) \, dW_s = f(t) W_t - \int_0^t \frac{df(s)}{ds} \cdot W_s \, ds$$

or

$$\int_0^t f(s) \, dW_s = f(t) W_t - \int_0^t W_s df_s.$$

integration by parts formula

This last formula is also known as the **integration by parts formula** for stochastic integrals.

Problem K. Use Itô's formula to show that for $n \geq 1$

$$d(W_t^{2n}) = n(2n-1) W_t^{2n-2} dt + 2n W_t^{2n-1} dW_t.$$

Solution. Let $Y_t = U(t, W_t) = W_t^{2n}$, $n \geq 1$. Then, use Itô's formula to get the desired result. In fact, $\frac{\partial U(t,W_t)}{\partial t} = 0$, $\frac{\partial U(t,W_t)}{\partial W_t} = 2n \cdot W_t^{2n-1}$, and $\frac{\partial^2 U(t,W_t)}{\partial (W_t)^2} = 2n \cdot (2n-1) \cdot W_t^{2n-2}$ all of which satisfy the requirements of Theorem 21.
This result can be used to compute moments for standard Brownian motion. Let $W_t \in \mathbf{R}$ and $W_0 = 0$. Define

$$\beta_{2n}(t) = E[W_t^{2n}], \quad n = 0, 1, 2, \ldots; t \geq 0.$$

Then, one can prove that

$$\beta_{2n}(t) = n \cdot (2n-1) \cdot \int_0^t \beta_{2n-2}(s) ds.$$

Since $\beta_0(t) = 1$ and $\beta_2(t) = t$, and so on, the formula can be used recursively to find higher moments, like $\beta_4(t) = 3t^2$, $\beta_6(t) = 15t^3$, etc...
The proof is based on the fact that

$$W_t^{2n} = n \cdot (2n-1) \int_0^t W_t^{2n-2} dt + 2n \int_0^t W_t^{2n-1} dW_t$$

and, since $W_t^{2n-2} \geq 0$, one can use Fubini's Theorem and interchange integration and expectation in the first integral on the right-hand side, to find

$$\beta_{2n}(t) \equiv E[W_t^{2n}] = n \cdot (2n-1) \int_0^t E[W_t^{2n-2}] dt + 2n \cdot E\Big[\int_0^t W_t^{2n-1} dW_t\Big].$$

Approximating the stochastic integral on the right-hand side by sums and proving that the sequence of sums converges in mean square to 0 completes the proof.

7. Vector Valued Itô Integrals

Let $\{\mathbf{W}_t : t \geq 0\}$ be an m-dimensional Wiener process with independent components (Figure 3) associated with an increasing family of σ-algebras $\{\mathcal{F}_t : t \geq 0\}$. That is, $\mathbf{W}_t = (W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(m)})$ where $W^{(j)}$, $j = 1, 2, \ldots, m$ are scalar Wiener processes with respect to $\{\mathcal{F}_t : t \geq 0\}$ which are pairwise independent. Thus, each $W_t^{(j)}$ is $\mathcal{F}_t$-measurable with

$$E[W_t^{(j)} \mid \mathcal{F}_0] = 0, \quad E[W_t^{(j)} - W_S^{(j)} \mid \mathcal{F}_s] = 0$$

w.p.1, for $0 \leq s \leq t$ and $j = 1, 2, \ldots, m$. In addition,

$$E[(W_t^{(i)} - W_s^{(i)})(W_t^{(j)} - W_s^{(j)})] = (t - s)\delta_{i,j}, \tag{1.40}$$

w,p.1, for $0 \leq s \leq t$ and $i, j = 1, 2, \ldots, m$ where $\delta_{i,j}$ is the Kronecker delta symbol. We shall consider d-dimensional vector functions $\mathbf{e} : [0, T] \times \Omega \mapsto \mathbf{R}^d$ with components $e^{(k)}$ satisfying $\sqrt{e^{(k)}} \in \mathcal{L}^2_\omega[0, T]$ or $\mathcal{L}^2[0, T]$ for $k = 1, 2, \ldots, d$ and $d \times m$-matrix functions $\mathbf{F} : [0, T] \times \Omega \mapsto \mathbf{R}^{d \times m}$ with components $F^{i,j} \in \mathcal{L}^2_\omega[0, T]$ or $\mathcal{L}^2[0, T]$ for $k = 1, 2, \ldots d$ and $j = 1, 2, \ldots, m$. In analogy with the scalar case we denote by $\mathbf{e}_t$ and $\mathbf{F}_t$ the vector and matrix valued random variables taken by $\mathbf{e}$ and $\mathbf{F}$ at an instant t. Then, we write symbolically as a d-dimensional vector stochastic differential

$$d\mathbf{X}_t = \mathbf{e}_t \, dt + \mathbf{F}_t \, d\mathbf{W}_t \tag{1.41}$$

the vector stochastic integral expression

$$\mathbf{X}_t - \mathbf{X}_s = \int_s^t \mathbf{e}_u \, du + \int_s^t \mathbf{F}_u \, d\mathbf{W}_u \tag{1.42}$$

for any $0 \leq s \leq t \leq T$, which we interpret componentwise as

$$X_t^{(k)} - X_s^{(k)} = \int_s^t e_u^{(k)} \, du + \sum_{j=1}^m \int_s^t F_u^{k,j} \, dW_u^{(j)}, \tag{1.43}$$

w.p.1, for $k = 1, 2, \ldots, d$. When $d = 1$ this convers the scalar case with several independent white noise processes. For a preassigned $\mathcal{F}_0$-measurable $\mathbf{X}_0$ the resulting d-dimensional stochastic process $\mathbf{X} = \{\mathbf{X}_t = (X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(d)}) : t \geq 0\}$ enjoys similar properties componentwise to those listed in the previous sections for scalar differentials involving a single Wiener process, with additional properties relating the different components.

The actual properties depend on whether the $\sqrt{e^{(k)}}$ and $F^{k,j}$ belong to $\mathcal{L}^2[0, T]$ or just to $\mathcal{L}^2_\omega[0, T]$. In the former case with $e \equiv 0$, for example, we have

$$E[X_t^{(k)} - X_s^{(k)} \mid \mathcal{F}_s] = 0$$

$$E[(X_t^{(k)} - X_s^{(k)})(X_t^{(i)} - X_s^{(i)}) \mid \mathcal{F}_s] = \sum_{j=1}^m \int_s^t E[F_u^{k,j} F_u^{i,j}] \, du, \tag{1.44}$$

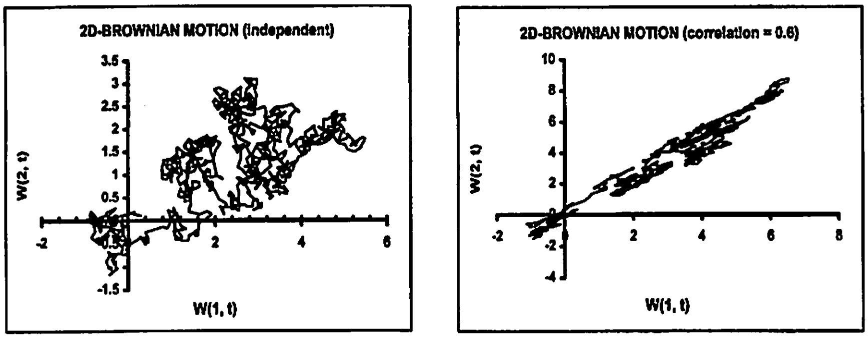

Figure 1.3: *Brownian motion in* $\mathbf{R}^2 : \mathbf{W}_t = (W_t^{(1)}, W_t^{(2)})$. *In the first case* $W_t^{(1)}$, *and* $W_t^{(2)}$ *are independent; in the second case their correlation is 0.6.*

w.p.1, for $0 \le s \le t \le T$ and $k, i = 1, 2, \ldots, d$. Here (1.44) follows form the independence of the components of $\mathbf{W}$ and the identity (1.40), which we could write symbolically as $E[dW_t^{(i)} \cdot dW_t^{(j)}] = \delta_{i,j} dt$. As in the scalar case this leads to additional terms in the chain rule formula for the transformation of the vector stochastic differential (1.41). We are now ready for the next important theorem.

Theorem 24. (Multi-dimensional Itô's Formula) *Let* $U : [0,T] \times \mathbf{R}^d \mapsto \mathbf{R}$ *have continuous partial derivatives* $\frac{\partial U}{\partial t}$, $\frac{\partial U}{\partial x_k}$, $\frac{\partial^2 U}{\partial x_k \partial x_i}$, $k, i = 1, 2, \ldots, d$, *and define a scalar process* $\{Y_t : t \ge 0\}$ *by*

$$Y_t = U(t, \mathbf{X}_t) = U(t, X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(d)}),$$

w.p.1, where $\mathbf{X}_t$ *satisfies* (1.41). *Then the stochastic differential for* Y_t *is given by*

$$\begin{aligned} dY_t = &\left[\frac{\partial U}{\partial t} + \sum_{k=1}^{d} e_t^{(k)} \cdot \frac{\partial U}{\partial x_k} + \frac{1}{2} \sum_{j=1}^{m} \sum_{i,k=1}^{d} F_t^{i,j} \cdot F_t^{k,j} \frac{\partial^2 U}{\partial x_i \partial x_k} \right] dt \\ &+ \sum_{j=1}^{m} \sum_{i}^{d} F_t^{i,j} \frac{\partial U}{\partial x_i} dW_t^{(j)}, \end{aligned} \tag{1.45}$$

where the partial derivatives are evalutated at $(t, \mathbf{X}_t)$. *The multi-component analogue of the Itô's formula is also written in vector-matrix notation as*

$$dY_t = \left[\frac{\partial U}{\partial t} + \mathbf{e}_t^T \nabla U + \frac{1}{2} \mathrm{tr}(\mathbf{F}_t \mathbf{F}_t^T \nabla^2 U) \right] dt + \nabla U^T \mathbf{F}_t \, d\mathbf{W}_t, \tag{1.46}$$

where ∇ *is the gradient operator,* ∇^2 *is the matrix of second order spatial partial derivatives of* U, ^{T} *the vector or matrix transpose operation and* tr *is the trace of the inscribed matrix.*

Proof Although the proof is cumbersome, it is just a straightforward modification of the proof provided for the scalar case.

Problem L. Let $X_t^{(1)}, X_t^{(2)}$ satisfy the scalar stochastic differentials

$$dX_t^{(i)} = e_t^{(i)} dt + f_t^{(i)} dW_t^{(i)}$$

for $i = 1, 2$ and let $U(t, X_t^{(1)}, X_t(2)) = X_t^{(1)} \cdot X_t(2)$. Find the stochastic differential for the product process $Y_t = X_t^{(1)} \cdot X_t^{(2)}$.

Solution. The answer depends on whether the Wiener processes $W_t^{(1)}$ and $W_t^{(2)}$ are independent (Figure 3) or dependent (Figure 4), so we will have to consider two cases. In the former case, (1.41) can be written as

$$d\begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \end{pmatrix} = \begin{pmatrix} e_t^{(1)} \\ e_t^{(2)} \end{pmatrix} dt + \begin{pmatrix} f_t^{(1)} & 0 \\ 0 & f_t^{(2)} \end{pmatrix} d\begin{pmatrix} W_t^{(1)} \\ W_t^{(2)} \end{pmatrix}.$$

Now, in our case, we find that

$$e_t = \begin{pmatrix} e_t^{(1)} \\ e_t^{(2)} \end{pmatrix}, \quad \frac{\partial U}{\partial t} = 0, \quad \nabla U = \begin{pmatrix} X_t^{(2)} \\ X_t^{(1)} \end{pmatrix},$$

$$F_t = \begin{pmatrix} f_t^{(1)} & 0 \\ 0 & f_t^{(2)} \end{pmatrix}, \quad \text{and} \quad \nabla^2 U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Applying formula (1.46) we finally find

$$dY_t = (e_t^{(1)} X_t^{(2)} + e_t^{(2)} X_t^{(1)})\, dt + f_t^{(1)} X_t^{(2)}\, dW_t^{(1)} + f_t^{(2)} X_t^{(1)}\, dW_t^{(2)}.$$

When $W_t^{(1)} = W_t^{(2)} = W_t$ the vector differential can be written as

$$d\begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \end{pmatrix} = \begin{pmatrix} e_t^{(1)} \\ e_t^{(2)} \end{pmatrix} dt + \begin{pmatrix} f_t^{(1)} \\ f_t^{(2)} \end{pmatrix} dW_t$$

and using again (1.46) with $F_t = \begin{pmatrix} f_t^{(1)} \\ f_t^{(2)} \end{pmatrix}$ gives

$$dY_t = (e_t^{(1)} X_t^{(2)} + e_t^{(2)} X_t^{(1)} + f_t^{(1)} f_t^{(2)})\, dt + (f_t^{(1)} X_t^{(2)} + f_t^{(2)} X_t^{(1)}) dW_t.$$

Problem M. Show that

$$d(W_t^{(1)} W_t^{(2)}) = W_t^{(2)}\, dW_t^{(1)} + W_t^{(1)}\, dW_t^{(2)}$$

for independent Wiener processes $W_t^{(1)}$ and $W_t^{(2)}$, whereas

$$d((W_t)^2) = dt + 2W_t\, dW_t$$

when $W_t^{(1)} = W_t^{(2)} = W_t$.

Solution. Use the results from the previous problem with $e_t^{(1)} = e_t^{(2)} = 0$ and $f_t^{(1)} = f_t^{(2)} = 1$.

If $\mathbf{W}_t = \{(W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(m)})\}$ is such that the scalar Brownian motions $W_t^{(i)}$, $i = 1, 2, \ldots, m$ are not independent and $\mathrm{Corr}(W_t^{(i)}, W_t^{(j)}) = \rho_{ij}$, $i, j = 1, 2, \ldots, m$ formula (1.45) is no longer valid and must be replaced by the following:

$$\begin{aligned} dY_t = &\left[\frac{\partial U}{\partial t} + \sum_{k=1}^{d} e_t^{(k)} \cdot \frac{\partial U}{\partial x_k} + \frac{1}{2}\sum_{j=1}^{m}\sum_{i,k=1}^{d} \rho_{i,k} F_t^{i,j} \cdot F_t^{k,j} \frac{\partial^2 U}{\partial x_i \partial x_k}\right] dt \\ &+ \sum_{j=1}^{m}\sum_{i}^{d} F_t^{i,j} \frac{\partial U}{\partial x_i} dW_t^{(j)}, \end{aligned} \tag{1.47}$$

8. Fisk-Stratonovich Integral

The Itô integral $\int_0^T f(t,\omega) dW_t(\omega)$ for an integrand $f \in \mathcal{L}^2[0,T]$ is equal to the mean-square limit of the sums

$$S_n(\omega) = \sum_{j=1}^{n} f(c_j^{(n)}, \omega) \cdot \Delta W_{t_{j+1}^{(n)}} \tag{1.48}$$

with evaluation points $c_i^{(n)} = t_j^{(n)}$ for partitions $0 < t_1^{(n)} < t_2^{(n)} < \ldots < t_{n+1}^{(n)} = T$ for which

$$\delta^{(n)} = \max_{1 \le j \le n} (t_{j+1}^{(n)} - t_j^{(n)}) \to 0$$

as $n \to \infty$. Other choices of evaluation points $t_j^{(n)} \le c_j^{(n)} \le t_{j+1}^{(n)}$ are possible, but generally lead to different random variables in the limit. While arbitrarily chosen evaluation points have little practical or theoretical use, those chosen systematically by

$$c_j^{(n)} = (1 - \lambda) t_j^{(n)} + \lambda t_{j+1}^{(n)} \tag{1.49}$$

for some fixed $0 \le \lambda \le 1$ lead to limits, which we shall denote here by

$$(\lambda) \int_0^T f(t,\omega) dW_t(\omega).$$

We note that the case $\lambda = 0$ is just the Itô integral. The other cases, $0 < \lambda \le 1$ differ in that the process they define with respect to a variable upper integration endpoint is in general no longer a martingale. Assuming that the integrand f

has continuously differentiable sample paths, we can apply Taylor's theorem as follows:

$$(1-\lambda)f(t_j^{(n)}) = (1-\lambda)f[(1-\lambda)t_j^{(n)} + \lambda t_{j+1}^{(n)}]$$
$$+ (1-\lambda)\cdot\lambda f'[(1-\lambda)t_j^{(n)} + \lambda t_{j+1}^{(n)}]\cdot(t_j^{(n)} - t_{j+1}^{(n)}) + O(|t_j^{(n)} - t_{j+1}^{(n)}|^2)$$

and

$$\lambda f(t_{j+1}^{(n)}) = \lambda f[(1-\lambda)t_j^{(n)} + \lambda t_{j+1}^{(n)}]$$
$$+ \lambda\cdot(1-\lambda)f'[(1-\lambda)t_j^{(n)} + \lambda t_{j+1}^{(n)}]\cdot(-t_j^{(n)} + t_{j+1}^{(n)}) + O(|t_j^{(n)} - t_{j+1}^{(n)}|^2).$$

Summing the last two expressions, we find that

$$f((1-\lambda)t_j^{(n)} + \lambda t_{j+1}^{(n)},\omega) = (1-\lambda)f(t_j^{(n)},\omega) + \lambda f(t_{j+1}^{(n)},\omega) + O(|t_j^{(n)} - t_{j+1}^{(n)}|).$$

Since the higher order terms do not contribute to the limit as $\delta^{(n)} \to 0$, we see that the (λ)-integrals could be evaluated alternatively as the mean-square limit of sums

$$\tilde{S}_n(\omega) = \sum_{j=1}^{n}[(1-\lambda)f[(1-\lambda)f(t_j^{(n)},\omega) + \lambda f(t_{j+1}^{(n)},\omega)]\cdot\Delta W_{t_{j+1}^{(n)}}. \tag{1.50}$$

In the general case the (λ)-integrals are usually defined in terms of the sums (1.50) rather than (1.48) with evaluation points (1.49), and we shall follow this practice here. To this purpose, we observe that for $f(t,\omega) = W_t(\omega)$ it is

$$(\lambda)\int_0^T W_t(\omega)dW_t(\omega) = \frac{1}{2}W_T(\omega)^2 + \left(\lambda - \frac{1}{2}\right)T. \tag{1.51}$$

This follows from the following mean-square limits

$$\sum_{j=1}^{n} W_{t_j^{(n)}}\cdot\Delta W_{t_{j+1}^{(n)}} \to \frac{1}{2}W_T^2 - \frac{1}{2}T$$

and

$$\sum_{j=1}^{n} W_{t_{j+1}^{(n)}}\cdot\Delta W_{t_{j+1}^{(n)}}$$
$$= \sum_{j=1}^{n}\left(W_{t_{j+1}^{(n)}} - W_{t_j^{(n)}}\right)^2 + \sum_{j=1}^{n} W_{t_j^{(n)}}\cdot\left(W_{t_{j+1}^{(n)}} - W_{t_j^{(n)}}\right)$$
$$\to T + \left(\frac{1}{2}W_T^2 - \frac{1}{2}T\right) = \frac{1}{2}(W_T^2 + T),$$

which are multiplied by $(1-\lambda)$ and λ, respectively, to give (1.51). Unlike any of the others, the symmetric case $\lambda + 1/2$ of the integral (1.51), which was introduced by Stratonovich, does not contain a term in addition to that given by classical calculus. It is now known as the **Fisk-Stratonovich integral** and denoted by

$$\int_0^T f_t \circ dW_t$$

for an integrand $f \in \mathcal{L}^2[0,T]$; it can be extended to integrands in $\mathcal{L}^2_\omega[0,T]$ in the same way as for Itô integrals. Usually, only the Itô and Fisk-Stratonovich integrals are widely used. As suggested by (1.51) the Fisk-Stratonovich integral obeys the transformations rules of classical calculus, and this is a major reason for its use. To see this, let $h : \mathbf{R} \mapsto \mathbf{R}$ be continuously differentiable and consider the Fisk-Stratonovich integral of $h(W_t)$. By the Taylor formula we have

$$h(W_{t^{(n)}_{j+1}}) = h(W_{t^{(n)}_j}) + h'(W_{t^{(n)}_j})\Delta W_{t^{(n)}_{j+1}} + \text{ higher order terms,}$$

so the sum (1.50) with $\lambda = 1/2$ is

$$\begin{aligned}\tilde{S}_n &= \sum_{j=1}^n h(W_{t^{(n)}_j}) \cdot \Delta W_{t^{(n)}_{j+1}} + \frac{1}{2}\sum_{j=1}^n h'(W_{t^{(n)}_j})\Delta W^2_{t^{(n)}_{j+1}} \\ &\quad + \text{ higher order terms} \\ &\overset{\mathcal{L}^2}{\to} \int_0^T h(W_t)dW_t + \frac{1}{2}\int_0^T h'(W_t)dt.\end{aligned}$$

Hence, we have show that there is a relationship between the Fisk-Stratonovich and the Itô integrals and it is given by:

$$\int_0^T h(W_t) \circ dW_t = \int_0^T h(W_t)dW_t + \frac{1}{2}\int_0^T h'(W_t)dt. \tag{1.52}$$

Now, let U be the antiderivative of h, so $U'(x) = h(x)$ and hence $U''(x) = h'(x)$. Applying Itô's formula to the transformation $Y_t = U(W_t)$, one obtains

$$U(W_T) - U(W_0) = \frac{1}{2}\int_0^T h'(W_t)dt + \int_0^T h(W_t)dW_t.$$

Thus, from (1.52) we see that the Fisk-Stratonovich integral satisfies

$$\int_0^T h(W_t) \circ dW_t = U(W_T) - U(W_0),$$

as in classical calculus. This result extends to more complicated stochastic differential equations. In fact, the interpretation of the (white-noise) equation

$$\frac{dX}{dt} = b(t, X_t) + \sigma(t, X_t) \cdot WN_t$$

is that X_t is the solution of the integral equation

$$X_t = X_0 + \int_0^t b(s, X_s)\,ds + \int_0^t \sigma(s, X_s)dW_s,$$

for some suitable interpretation of the stochastic integral in the right-hand side of the previous equality where Itô or Fisk-Stratonovich interpretations are just two of several reasonable choices. The question is: which interpretation gives the right mathematical model for the white-noise equation? The answer essentially depends on the specific application. Nevertheless, if

$$X_t = X_0 + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s) \circ dW_s$$

then, X_t is the solution of the following modified Itô equation (see e.g. Kloeden and Platen (1992) pp. 154-160):

$$X_t = X_0 + \int_0^t \left[b(s, X_s) + \frac{1}{2}\sigma'(s, X_s)\sigma(x, X_s) \right] ds + \int_0^t \sigma(s, X_s)\, dW_s \quad (1.53)$$

where σ' denotes the derivative of $\sigma(t, x)$ with respect to x. Clearly, when $\sigma(t, X_t)$ does not depend on X_t the two interpretations lead to the same solution. From this result, one gets also the relationship between the two stochastic integrals, i.e.:

$$\int_0^T \sigma(s, X_s) \circ dW_s = \int_0^T \frac{1}{2}\sigma'(s, X_s) \cdot \sigma(s, X_s)\, ds + \int_0^T \sigma(s, X_s)\, dW_s.$$

Problem N. Transform the following Stratonovich stochastic differential equations into Itô stochastic differential equations:

(a) $X_t = \gamma X_t\, dt + \alpha X_t \circ dW_t$;

(b) $X_t = \sin X_t \cos X_t\, dt + (t^2 + \cos X_t) \circ dW_t$.

Solution. Use (1.53) where $\sigma(s, X_s) = \alpha X_t$ for (a) and $t^2 + \cos X_t$ for (b). This gives

(a) $dX_t = \left(\gamma + \frac{1}{2}\alpha^2\right) dt + \alpha X_t\, dW_t$;

(b) $dX_t = \left(\sin X_t \cos X_t + \frac{1}{2}\sin X_t (t^2 + \cos X_t)\right) dt + (t^2 + \cos X_t)\, dW_t$;
which can be rewritten as
$dX_t = \frac{1}{2}\sin X_t [\cos X_t - t^2]\, dt + (t^2 + \cos X_t)\, dW_t$.

Problem O. Transform the following Itô stochastic differential equations into Stratonovich stochastic differential equations:

(a) $dX_t = rX_t\, dt + \alpha X_t\, dW_t$;

(b) $dX_t = 2e^{-X_t}\, dt + X_t^2\, dW_t$.

Solution. From (1.53) we know that any Stratonovich differential equation as

$$dX_t = rX_t\,dt + \alpha X_t \circ dW_t$$

produces an Itô differential equation like

$$dX_t = (r + \alpha^2/2)X_t\,dt + \alpha X_t\,dW_t$$

so that it is easy to guess that the Stratonovich differential equation for (a) is given by

$$dX_t = (r - \alpha^2/2)X_t\,dt + \alpha X_t \circ dW_t.$$

For (b), it is enough to solve with respect to $b(t, X_t)$the following

$$b(t, X_t) + (1/2)\sigma(t, X_t)\cdot\sigma'(t, X_t) = 2e^{-X_t)}$$

given that $\sigma(t, X_t) = X_t^2$. It is easily seen that the solution is provided by $b(t, X_t) = 2e^{-X_t)} - X_t^3$. Hence, the Stratonovich differential equation is

$$dX_t = (2e^{-X_t)} - X_t^3)\,dt + X_t^2 \circ dW_t.$$

Problem P. Suppose that $f \in \mathcal{L}^2[0, T]$ and is such that there exists $K < \infty$ and $\epsilon > 0$ so that

$$E[|f(s,\cdot) - f(t,\cdot)|^2] \le K|s-t|^{1+\epsilon};\ \ 0 \le s, t \le T.$$

In addition, assume that for every partition $0 = t_1^{(n)} < t_2^{(n)} < \dots, < t_{n+1}^{(n)} = T$ we have that $\max_{1\le j\le n} t_{j+1}^{(n)} - t_j^{(n)} \le n^{-1}$. Prove that

$$\int_0^T f(t,\omega)dW_t = \int_0^T f(t,\omega)\circ dW_t \ \text{ in } L^1(P^0)$$

where P_0 is the probability law of W_t starting at 0.

Solution. Letting $t_j^{(n)\star} = (1/2)\cdot(t_{j+1}^{(n)} + t_j^{(n)})$, $\delta_n = \max_{1\le j\le n} t_{j+1}^{(n)} - t_j^{(n)}$, and following the definition of the two stochastic integrals, one finds that

$$E[|S_n - \tilde{S}_n|] \le \sum_{j=1}^{n} E[|f(t_j^{(n)},\omega) - f(t_j^{(n)\star},\omega)|\cdot|\Delta W(t_{j+1}^{(n)},\omega)|$$

and, by the Cauchy-Schwarz inequality and the assumption of the text,

$$\le \sum_{j=1}^{n}\sqrt{E[|f(t_j^{(n)},\omega) - f(t_j^{(n)\star})|^2]}\cdot\sqrt{E[\Delta W(t_{j+1}^{(n)},\omega)]^2}$$

$$\le \sum_{j=1}^{n}\sqrt{K|t_j^{(n)} - t_j^{(n)\star}|^{1+\epsilon}}\cdot\sqrt{|t_j^{(n)} - t_j^{(n)\star}|} \le K^{1/2}\cdot\sum_{j=1}^{n}\delta_n^{1+\epsilon/2}.$$

Since, by assumption, we have that $\delta_n \le n^{-1}$ for all $n = 1, 2, \dots$ we find that

$$\lim_n E[|S_n - \tilde{S}_n|] = 0$$

which proves the assertion.

9. Stochastic Integrals with Stopping Times

If τ is an $\{\mathcal{F}_t\}$-stopping time, then $I_{[0,\tau]\times\Omega}(t,\omega)$ is a bounded continuous $\mathcal{F}_t$-adapted process. In fact, boundedness and right-continuity are obvious, while $\{\mathcal{F}_t\}$-measurability follows from being for each $t \geq 0$

$$\{\omega : I_{[0,\tau]\times\Omega}(t,\omega) \leq r\} = \begin{cases} \emptyset & \text{if } r < 0, \\ \{\omega : \tau(\omega) < t\} & \text{if } 0 \leq r < 1, \\ \Omega & \text{if } r \geq 1, \end{cases}$$

that is $I_{[0,\tau]\times\Omega}(t,\omega)$ is $\{\mathcal{F}_t\}$-measurable. It follows that $\{I_{[0,\tau]\times\Omega}(t,\omega) : t \in \mathbf{R}_+\}$ is also predictable (a continuous adapted process is predictable) and, hence, the process of defining a stochastic integral with stopping times is well posed.

Definition 25. *Let $f \in \mathcal{L}^2_T$ and let τ be an $\{\mathcal{F}_t\}$-stopping time such that $0 \leq \tau \leq T$. Then, $\{I_{[0,\tau]\times\Omega}(t,\omega) \cdot f(t,\omega) : t \in [0,T]\}$ is also in $\mathcal{L}^2_T$ and we define*

$$\int_0^\tau f(s,\omega)\, dW_s(\omega) = \int_0^T I_{[0,s]\times\Omega}(t,\omega) \cdot f(s,\omega)\, dW_s(\omega). \tag{1.54}$$

Furthermore, if ρ is another stopping time with $0 \leq \rho \leq \tau$, we define

$$\int_\rho^\tau f(s,\omega)\, dW_s(\omega) = \int_0^\tau f(s,\omega)\, dW_s(\omega) - \int_0^\rho f(s,\omega)\, dW_s(\omega). \tag{1.55}$$

It is easy to check that

$$\int_\rho^\tau f(s,\omega)\, dW_s(\omega) = \int_0^T I_{(\rho,\tau]\times\Omega}(t,\omega) \cdot f(s,\omega)\, dW_s(\omega).$$

Theorem 26. *Let $f \in \mathcal{L}_2[a,b]$ and ξ is a real-valued bounded $\mathcal{F}_a$-measurable random variable. Then, $xi \cdot f \in \mathcal{L}_2[a,b]$ and*

$$\int_a^b \xi \cdot f(t,\omega)\, dW_t(\omega) = \xi \cdot \int_a^b f(t,\omega)\, dW_t(\omega). \tag{1.56}$$

Proof. It is clear that $\xi \cdot f \in \mathcal{L}^2[a,b]$. In addition, if f is a simple process then (1.56) follows from the definition of stochastic integral. For general $f \in \mathcal{L}^2[a,b]$, let $\{\phi_n : n \in \mathbf{N}\}$ be a sequence of simple processes satisfying the requirement

$$\lim_n E \int_a^b |f(t,\omega) - \phi_n(t,\omega)|^2\, dt = 0.$$

Applying (1.56) to each ϕ_n and taking $n \to \infty$ the assertion follows.

Next, we state another theorem which we have partially already seen

Theorem 27. *Let $f \in \mathcal{L}^2[a,b]$. Then,*

$$E\left(\int_a^b f(t,\omega)\,dW_t(\omega) \mid \mathcal{F}_a\right) = 0 \tag{1.57}$$

$$E\left(|\int_a^b f(t,\omega)\,dW_t(\omega)|^2 | \mathcal{F}_a\right) = E\left(\int_a^b |f(t,\omega)|^2\,dt \mid \mathcal{F}_a\right) = \int_a^b E\left(|f(t,\omega)|^2 | \mathcal{F}_a\right) dt. \tag{1.58}$$

Proof. By the definition of conditional expectation, (1.57) holds iff

$$E\left(I_A \cdot \int_a^b f(t,\omega)\,dW_t(\omega)\right) = 0$$

for all sets $A \in \mathcal{F}_a$. By Theorem 26 and Theorem 6.b ,

$$E\left(I_A \cdot \int_a^b f(t,\omega)\,dW_t(\omega)\right) = E\left(\int_a^b I_A \cdot f(t,\omega)\,dW_t(\omega)\right) = 0$$

as required. The proof of (1.58) is similar. In fact, for any $A \in \mathcal{F}_a$ we have

$$E\left(|\int_a^b f(t,\omega)\,dW_t(\omega)|^2 | \mathcal{F}_a\right) = E\left(I_A \cdot |\int_a^b f(t,\omega)\,dW_t(\omega)|^2\right)$$

by Theorem 26

$$= E\left(|\int_a^b I_A \cdot f(t,\omega)\,dW_t(\omega)|^2\right)$$

which, by Theorem 6.c, can be rewritten as

$$\begin{aligned}
&= E\left(\int_a^b |I_A \cdot f(t,\omega)|^2\,dt\right) \\
&= \int_a^b E\left(I_A |f(t,\omega)|^2\right) dt \\
&= \int_a^b E|f(t,\omega)|^2 | \mathcal{F}_a)\,dt
\end{aligned}$$

It is not difficult to show that Theorem 6 can be extended as follows

Theorem 28. *Let $f \in \mathcal{L}^2_T$ and let ρ, τ be two stopping times such that $0 \le \rho \le \tau \le T$. Then*

$$E[\int_\rho^\tau f(s,\omega)\,dW_s(\omega) = 0, \quad E|\int \rho^\tau f(s,\omega)\,dW_s(\omega)|^2 = E[\int_\rho^\tau |f(s,\omega)|^2\,dt.$$

The last theorem can be improved to generalize Theorem 27.

Theorem 29. *Let $f \in \mathcal{L}_T^2$ and let ρ, τ be two stopping times such that $0 \le \rho \le \tau \le T$. Then*

$$E\left(\int_\rho^\tau f(s,\omega)\, dW_s(\omega) \mid \mathcal{F}_\rho\right) = 0, \tag{1.59}$$

$$E\left(\left|\int_\rho^\tau f(s,\omega)\, dW_s(\omega)\right|^2 \mid \mathcal{F}_\rho\right) = E\left(\int_\rho^\tau |f(s,\omega)|^2\, ds \mid \mathcal{F}_\rho\right). \tag{1.60}$$

Proof. Before proving the last statement, we need two useful lemmas.

Lemma 1. *Let $f \in \mathcal{L}_T^2$ and let τ be a stopping time such that $0 \le \tau \le T$. Then,*

$$\int_0^\tau f(s,\omega)\, dW_s(\omega) = I(\tau).$$

Before stating and proving the second lemma, we need to introduce a few definition first. A stochastic process $X = \{X_t : t \in \mathbf{R}_+\}$ is called **square-integrable** if $E[|X_t|^2] < \infty$ for every $t \in \mathbf{R}_+$. If $M = \{M_t : t \in \mathbf{R}_+\}$ is a real-valued square-integrable continuous martingale, then there exists a unique continuous integrable adapted increasing process denoted by $\{\langle M, M\rangle_t\}$ such that $\{M_t^2 - \langle M, M\rangle_t\}$ is a continuous martingale vanishing at $t = 0$. The process $\{\langle M, M\rangle_t\}$ is called the **quadratic variation** of M. Now,

square-integrable

quadratic variation

Lemma 2. *Let $f \in \mathcal{L}_T^2$. Then the indefinite integral $I = \{I(t) : t \in \mathbf{R}_+\}$ is a square integrable continuous martingale and its quadratic variation is given by*

$$\langle I, I\rangle_t = \int_0^t |f(s,\omega)|^2\, ds, \quad 0 \le t \le T. \tag{1.61}$$

Proof. From Theorem 10 we already know that $I = \{I(t) : t \in \mathbf{R}_+\}$ is a square integrable continuous martingale. To establish (1.61) we have to use the definition of quadratic variation and show that $\{I^2(t) - \langle I, I\rangle_t\}$ is a continuous martingale vanishing at $t = 0$. But, obviously, $I^2(0) - \langle I, I\rangle_0 = 0$. Moreover, if $0 \le r < t \le T$, by Theorem 27,

$$\begin{aligned} E[I^2(t) - \langle I, I\rangle_t \mid \mathcal{F}_r] &= I^2(r) - \langle I, I\rangle_r + 2I(r)\cdot E\left[\int_r^t f(s,\omega)\, dW_s(\omega) \mid \mathcal{F}_r\right] \\ &\quad + E\left[\left|\int_r^t f(s,\omega)\, dW_s(\omega)\right|^2 \mid \mathcal{F}_r\right] - E\left[\int_r^t |f(s,\omega)|^2\, ds \mid \mathcal{F}_r\right] \\ &= I^2(r) - \langle I, I\rangle_r \end{aligned}$$

as desired.

To prove Theorem 29 we notice that by Lemma 2 and the Doob martingale stopping theorem it is

$$E[I(\tau) \mid \mathcal{F}_\rho] = I(\rho)$$

and

$$E[I^2(\tau) - \langle I, I\rangle_\tau \mid \mathcal{F}_\rho] = I^2(\rho) - \langle I, I\rangle_\rho.$$

Applying Lemma 1 we see from the first of the two equalities just stated that

$$E\left[\int_\rho^\tau f(s,\omega)\, dW_s(\omega) \mid \mathcal{F}_\rho\right] = E[I(\tau) - I(\rho) \mid \mathcal{F}_\rho] = 0.$$

Also, by the same equalities above, we find

$$\begin{aligned} E[|I(\tau) - I(\rho)|^2 | \mathcal{F}_\rho] &= E[I^2(\tau) \mid \mathcal{F}_\rho] - 2 \cdot I(\rho) \cdot E[I(\tau) \mid \mathcal{F}_\rho] + I^2(\rho) \\ &= E[I^2(\tau) \mid \mathcal{F}_\rho] - I^2(\rho) = E[\langle I, I\rangle_\tau - \langle I, I\rangle_\rho \mid \mathcal{F}_\rho] \\ &= E\left[\int_\rho^\tau |f(s,\omega)|^2\, ds \mid \mathcal{F}_\rho\right] \end{aligned}$$

which by Lemma 1 is the required (1.58).

Problem P. Let $f, g \in \mathcal{L}_T^2$ and let ρ, τ be two stopping times such that $0 \leq \rho \leq \tau \leq T$. Show that

$$E\left[\int_\rho^\tau f(s,\omega)\, dW_s(\omega) \cdot \int_\rho^\tau g(s,\omega)\, dW_s(\omega)\, ds \mid \mathcal{F}_\rho\right] = E\left[\int_\rho^\tau f(s,\omega) \cdot g(s,\omega)\, ds \mid \mathcal{F}_\rho\right].$$

Solution. Using Theorem 29, one has

$$\begin{aligned} 4E&\left[\int_\rho^\tau f(s,\omega)\, dW_s(\omega) \cdot \int_\rho^\tau g(s,\omega)\, dW_s(\omega)\, ds \mid \mathcal{F}_\rho\right] \\ &= E\left[\left|\int_\rho^\tau [f(s,\omega) + g(s,\omega)]\, dW_s(\omega)\right|^2 \mid \mathcal{F}_\rho\right] \\ &\qquad - E\left[\left|\int_\rho^\tau [f(s,\omega) - g(s,\omega)]\, dW_s(\omega)\right|^2 \mid \mathcal{F}_\rho\right] \\ &= E\left[\int_\rho^\tau [f(s,\omega) + g(s,\omega)]^2\, ds \mid \mathcal{F}_\rho\right] - E\left[\int_\rho^\tau [f(s,\omega) - g(s,\omega)]^2\, ds \mid \mathcal{F}_\rho\right] \\ &= 4E\left[\int_\rho^\tau f(s,\omega) \cdot g(s,\omega)\, ds \mid \mathcal{F}_\rho\right] \end{aligned}$$

which completes the proof of the statement.

The results just seen can be extended to the multidimensional case. The interested reader is referred to the book by Mao (pp. 27-30).

10. Moment Inequalities, Gronwall-Type Inequalities

In this section Itô formula is applied to establish several important moment inequalities for stochastic integrals as well as the exponential martingale inequality. In this section, we let $W_t(\omega) = (W_{1t}(\omega), \dots, W_{mt}(\omega))^T$, $t \in \mathbf{R}_+$ be an m-dimensional Brownian motion defined on a complete probability space $(\Omega, \mathcal{F}, P)$ adapted to the filtration $\{\mathcal{F}_t : t \in \mathbf{R}_+\}$.

Theorem 30. *Let $p \geq 2$. Let g be an $\mathbf{R}^{d\times m}$-valued function in $\mathcal{L}^2_T$ such that*

$$E[\int_0^T |g(s,\omega)|^p\, ds < \infty.$$

Then

$$E\Big[|\int_0^T g(s,\omega)\, dW_s(\omega)|^p\Big] \leq \Big(\frac{p(p-1)}{2}\Big)^{p/2} \cdot T^{\frac{p-2}{2}} \cdot E\Big[\int_0^T |g(s,\omega)|^p\, ds\Big]. \quad (1.62)$$

In particular, for $p = 2$ there is equality.

Proof. For $p = 2$ the required result follows from Theorem 29. When $p > 2$ for $0 \leq t \leq T$, letting $x_0(\omega) = 0$ for any $\omega \in \Omega$, one can set

$$x_t(\omega) - x_0(\omega) = \int_0^t g(s,\omega)\, dW_s(\omega)$$

or, equivalently, using an Itô process representation,

$$dx_t(\omega) = 0\, dt + g(t,\omega)\, dW_t(\omega).$$

Now, if we write $|x_t(\omega)|$ for $tr[x_t(\omega)^T x_t(\omega)]^{1/2}$, using the multi-dimensional Itô formula, one finds that $|x_t(\omega)|^p$ can be written as an Itô process as well:

$$\begin{aligned} d|x_t(\omega)|^p &= \frac{1}{2} \cdot tr[g_t(\omega) \cdot g_t(\omega)^T \cdot p(p-1)|x_t(\omega)|^p]\, dt + [p \cdot |x_t(\omega)|^{p-1}]^T \cdot g_t(\omega)\, dW_t \\ &= \frac{p(p-1)}{2} \cdot |x_t(\omega)|^{p-2} \cdot |g_t(\omega)|^2\, dt + p|x_t(\omega)|^{p-1} \cdot g_t(\omega)\, dW_t(\omega). \end{aligned}$$

Integrating both sides, taking expectations and using the multidimensional analog of Theorem 6.b gives

$$E[|x_t(\omega)|^p] = \frac{p(p-1)}{2} \cdot E\Big[\int_0^t |x_s(\omega)|^{p-2} \cdot |g_s(\omega)|^2\, ds\Big].$$

Using Hölder's inequality

$$\begin{aligned} E[|x_t(\omega)|^p] &\leq \frac{p(p-1)}{2} \Big(E\Big[\int_0^t |x_s(\omega)|^p\, ds\Big]\Big)^{\frac{p-2}{p}} \cdot \Big(E\Big[\int_0^t |g_s(\omega)|^p\, ds\Big]\Big)^{\frac{2}{p}} \\ &= \frac{p(p-1)}{2} \Big(\int_0^t E[|x_s(\omega)|^p]\, ds\Big)^{\frac{p-2}{p}} \cdot \Big(\int_0^t E[|g_s(\omega)|^p]\, ds\Big)^{\frac{2}{p}} \end{aligned}$$

From a previous inequality it is easily seen that $E[|x_t(\omega)|^p]$ is nondecreasing in t. After simple algebraic manipulations this yields

$$E[|x_t(\omega)|^p] \leq \frac{p(p-1)}{2} \cdot \left[tE[|x_t(\omega)|^p]\right]^{\frac{p-2}{2}} \cdot \left(E\Big[\int_0^t |g_s(\omega)|^p\, ds\Big]\right)^{\frac{2}{p}}$$
$$= \left(\frac{p(p-1)}{2}\right)^{\frac{p}{2}} \cdot t^{\frac{p-2}{2}} \cdot E\Big[\int_0^t |g_s(\omega)|^p\, ds\Big]$$

from which the desired result follows once t is replaced by T.

Theorem 31. *Under the same assumptions as Theorem 30,*

$$E\left(\sup_{0\leq t\leq T} \Big|\int_0^t g_s(\omega)\, dW_s(\omega)\Big|^p\right) \leq \left(\frac{p^3}{p-1}\right)^{p/2} \cdot T^{\frac{p-2}{p}} \cdot E\Big[\int_0^T |g_s(\omega)|^p\, ds.$$

Proof. Since $\int_0^t g_s(\omega)\, dW_S(\omega)$ is an $\mathbf{R}^d$-valued martingale, one can use Doob Martingale Inequality and write

$$E\left(\sup_{0\leq t\leq T} \Big|\int_0^t g_s(\omega)\, dW_s(\omega)\Big|^p\right) \leq \left(\frac{p}{p-1}\right)^p \cdot E\Big[\Big|\int_0^T g_s(\omega)\, dW_s(\omega)\Big|^p.$$

A simple application of Theorem 30 completes the proof of the Theorem.

Theorem 32. (Burkholder-Davis-Gundy inequality) *Let g be an $\mathbf{R}^{d\times m}$-valued measurable $\mathcal{F}_t$-adapted process such that $\int_0^T |g_s(\omega)|^2\, ds < \infty$ a.s. for every $T > 0$. Define $t > 0$,*

$$x_t(\omega) = \int_0^t g_s(\omega)\, dW_s(\omega), \quad \text{and} \quad A_t(\omega) = \int_0^t |g_s(\omega)|^2\, ds.$$

Then, for every $p > 0$, there exist universal positive constants c_p, C_p (depending only on p), such that

$$c_p \cdot E[|A_t(\omega)|]^{p/2} \leq E\left(\sup_{0\leq s\leq t} |x_s(\omega)|^p\right) \leq C_p \cdot E[|A_t(\omega)|]^{p/2}.$$

Proof. See Mao (1997), pp. 40-3.

Theorem 33. (Gronwall's inequality) *Let $T > 0$ and $c \geq 0$. Let $u(\cdot)$ be a Borel measurable bounded nonnegative function on $[0,T]$, and let $v(\cdot)$ be a nonnegative integrable function on $[0,T]$. If*

$$u(t) \leq c \cdot \int_0^t v(s) \cdot u(s)\, ds \;\; \forall t \in [0,T],$$

then

$$u(t) \leq \exp\left\{\int_0^t v(s)\, ds\right\} \;\; \forall t \in [0,T].$$

Proof. Without loss of generality, it is possible to assume that $c > 0$. Set

$$z(t) = c + \int_0^t v(s) \cdot u(s)\, ds$$

for $0 \le t \le T$. This clearly implies that $z(0) = c$. In addition, one has also

$$\frac{z'(t)}{z(t)} = \frac{v(t) \cdot u(t)}{z(t)}$$

and integrating both sides with respect to t between 0 and t gives

$$\ln(z(s)) \,|_0^t = \int_0^t \frac{v(s) \cdot u(s)}{z(s)}\, ds$$

or

$$\ln(z(t)) = \ln(c) + \int_0^t \frac{v(s) \cdot u(s)}{z(s)}\, ds.$$

By our definition of $z(\cdot)$, $u(s) \le z(s)$ for $0 \le s \le T$ and, therefore,

$$\ln(z(t)) \le \ln(c) + \int_0^t v(s)\, ds$$

which implies

$$u(t) \le z(t) \le c \cdot \int_0^t v(s) \cdot u(s)\, ds$$

for $0 \le t \le T$ as we were supposed to prove.

Theorem 34. (Bihari's inequality) *Let $T > 0$ and $c > 0$. Let $K : \mathbf{R}_+ \mapsto \mathbf{R}_+$ be a continuous nondecreasing function such that $K(t) > 0$ for all $t > 0$. Let $u(\cdot)$ be a Borel measurable bounded nonnegative function on $[0, T]$, and let $v(\cdot)$ be a nonnegative integrable function on $[0, T]$. If*

$$u(t) \le c + \int_0^t v(s) K(u(s))\, ds \quad \forall t \in [0, T],$$

then

$$u(t) \le G^{-1}\left(G(c) + \int_0^t v(s)\, ds\right)$$

holds for all such $t \in [0, T]$ that

$$G(c) + \int_0^t v(s)\, ds \in Dom(G^{-1}),$$

and

$$G(r) = \int_0^r \frac{ds}{K(s)} \quad \text{on } r > 0,$$

and G^{-1} is the inverse function of G.

Proof. For $0 \le t \le T$, set

$$z(t) = c + \int_0^t v(s) \cdot K(u(s))\, ds.$$

It is easily seen that $u(t) \le z(t)$. Let also

$$H(t) \equiv G(z(t)).$$

Then, using the chain rule of classical calculus,

$$H'(t) = G'(z(t)) \cdot z'(t) = \frac{v(t) \cdot K(u(t))}{K(z(t))}.$$

Hence,

$$\int_0^t dH'(s) = \int_0^t \frac{v(s) \cdot K(u(s))}{K(z(s))}\, ds + H(0)$$

and, since $H(t) = G(z(t)), H(0) = G(z(0)) = G(c), u(s) \le z(s)$ for all $0 \le s \le T$, we have

$$G(z(t)) = G(c) + \frac{v(s) \cdot K(u(s))}{K(z(s))}\, ds \le G(c) + \int_0^t v(s)\, ds$$

for all $t \in [0,T]$ and, therefore, for all such $t \in [0,T]$ for which $G(c) + \int_0^t v(s)\, ds \in Dom(G^{-1})$, we have found

$$z(t) \le G^{-1}\left(G(c) + \int_0^t v(s)\, ds\right)$$

as we were supposed to.

Theorem 35 *Let $T > 0$, $\alpha \in [0,1)$ and $c > 0$. Let $u(\cdot)$ be a Borel measurable bounded nonnegative function on $[0,T]$, and let $v(\cdot)$ be a nonnegative integrable function on $[0,T]$. If*

$$u(t) \le c + \int_0^t v(s)[u(s)]^\alpha\, ds$$

for all $0 \le t \le T$, then

$$u(t) \le \left(c^{1-\alpha} + (1-\alpha)\int_0^t v(s)\, ds\right)^{\frac{1}{1-\alpha}}$$

holds for all $t \in [0,T]$.

Proof. For $t \in [0,T]$, set

$$z(t) = c + \int_0^t v(s)[u(s)]^\alpha\, ds.$$

Then, $u(t) \leq z(t)$ and $z(t) > 0$. Letting $H(t) \equiv [z(t)]^{1-\alpha}$ it is easily checked that

$$H'(s) = (1-\alpha) \cdot \frac{v(s)[u(s)]^{\alpha}}{[z(s)]^{\alpha}}$$

and, integrating over $[0, t]$ with respect to s,

$$[z(t)]^{1-\alpha} = c^{1-\alpha} + (1-\alpha) \cdot \int_0^t \frac{v(s)[u(s)]^{\alpha}}{[z(s)]^{\alpha}} \, ds \leq c^{1-\alpha} + (1-\alpha) \cdot \int_0^t v(s) \, ds$$

for all $t \in [0, T]$ from which the desired inequality follows immediately.

11. The Martingale Representation Theorem

Earlier it was shown that an Itô integral

$$X_t(\omega) = X_0(\omega) + \int_0^t f(t, \omega) \, dW_s$$

for any $f(\cdot, \cdot) \in \mathcal{L}^2[0, \infty]$ and $t \geq 0$ is always a martingale with respect to the filtration $\mathcal{F}_t$ generated by $\{W_s(\cdot) : s \leq t\}$ and with respect to the probability measure P^0. In this section we will prove that the converse is also true: i.e. any $\mathcal{F}_t$-martingale (with respect to P^0 can be represented as an Itô integral. This result is know as the **martingale representation theorem** and is important for many applications, for example in mathematical finance. Before we can state and prove the theorem, we need a couple of auxiliary lemmas.

Lemma 36. *Fix $T > 0$. The set of random variables*

$$\{\phi(W_{t_1}(\omega), \ldots, W_{t_n}(\omega)) : t_i \in [0, T], \phi \in C_0^{\infty}(\mathbf{R}^n); n \in \mathbf{N}\}$$

is dense in $L^2(\mathcal{F}_t, P^0)$, the space of squared integrable functions with respecto to P^0 and adapted to $\{\mathcal{F}_t : t \in \mathbf{R}_+\}$.

Proof. Let $\{t_i; i \in \mathbf{N}\}$ be a dense subset of $[0, T]$ and for each $n = 1, 2, \ldots$, let $\mathcal{H}_n$ be the σ-algebra generated by $W_{t_1}(\omega), \ldots, W_{t_n}(\omega)$. It is easily seen that $\mathcal{H}_n \subset \mathcal{H}_{n+1}$ and $\mathcal{F}_T$ is the smallest σ-algebra containing all the $\mathcal{H}_n$'s. Choose $g \in L^2(\mathcal{F}_t, P^0)$. By the martingale convergence theorem, one finds that

$$g = E[g \mid \mathcal{F}_T] = \lim_n E[g \mid \mathcal{H}_n].$$

The limit is pointwise a.e.(P^0). By the Doob-Dynkin Lemma (see Lemma B in the proof of Theorem 2) it is always possible to write

$$E[g \mid \mathcal{H}_n] = g_n(W_{t_1}, \ldots, W_{t_n})$$

for some Borel measurable function $g_n : \mathbf{R}^n \mapsto \mathbf{R}$, for each $n = 1, 2, \ldots$ Each such $g_n(W_{t_1}, \ldots, W_{t_n})$ can be approximated in $L^2(\mathcal{F}_t, P^0)$ by functions

$\phi_n(W_{t_1}, \dots, W_{t_n})$ where $\phi_n \in C_0^\infty(\mathbf{R}^n)$ (Lemma C in the proof of Theorem 2.) and the result follows.

Lemma 37. *The linear span (i.e. the set of linear combinations) of random variables of the type (Doleans exponential)*

$$\exp\left\{\int_0^T h(t)\, dW_t(\omega) - \frac{1}{2}\int_0^T h^2(t)\, dt\right\} \tag{1.63}$$

$h \in L^2[0,T]$, *deterministic, is dense in* $L^2(\mathcal{F}_t, P^0)$.

Proof. The idea is to prove that the set of stochastic integrals of the type (1.63) over the interval $[0,T]$ exhausts the space $L^2(\mathcal{F}_t, P^0)$. One way to do it is to show that the only $\mathcal{F}_t$-measurable, mean zero square integrable random variable which is orthogonal to all the stochastic integrals (1.63) is the zero-function. Suppose $g \in L^2(\mathcal{F}_t, P^0)$ is orthogonal to all functions. of the form (1.63). Then, if we let $h(t) = \sum_{i=1}^n \mu_i \cdot I_{(t_{i-1},t_i]}(t)$ for $t \in [0,T]$, with $\{\mu_i : i = 1,2,\dots n\}$ real constants, a simple application of the integration by parts rule gives

$$\begin{aligned}
\int_0^T h(t)\, dW_t(\omega) &= \sum_{i=1}^n \mu_i \cdot (W_{t_i}(\omega) - W_{t_{i-1}}(\omega)) \\
&= \mu_1(W_{t_1}(\omega) - W_{t_0}(\omega)) + (\mu_2 - \mu_1)(W_{t_2}(\omega) - W_{t_1}(\omega) + \dots \\
&+ (\mu_n - \mu_{n-1})(W_{t_n}(\omega) - W_{t_{n-1}}(\omega)) \\
&= \sum_{i=1}^n \lambda_i (W_{t_i}(\omega) - W_{t_{i-1}}(\omega))
\end{aligned}$$

where $\lambda_i = \mu_i - \mu_{i-1}$, $i = 1,2,\dots,n$. With this choice for h and letting $\Delta W_{t_i} = W_{t_i} - W_{t_{i-1}}$, $i = 1,2,\dots,n$ we have

$$G(\lambda_1, \dots, \lambda_n) \equiv \int_\Omega \exp\{\lambda_1 \Delta W_{t_1}(\omega) + \dots + \lambda_n \Delta W_{t_n}(\omega)\} \cdot g(\omega)\, dP^0(\omega) = 0 \tag{1.64}$$

for all $\lambda_1, \dots, \lambda_n$ and all $t_1, \dots, t_n \in [0,T]$. The function $G(\lambda_1, \dots, \lambda_n)$ is analytic in $\lambda \in \mathbf{R}^n$ and hence it has an analytic extension to the complex space $\mathbf{C}^n$ given by

$$G(z_1, \dots, z_n) = \int_\Omega \exp\{z_1 \Delta W_{t_1}(\omega) + \dots + z_n \Delta W_{t_n}(\omega)\} \cdot g(\omega)\, dP^0(\omega) \tag{1.65}$$

for all $(z_1, z_2, \dots, z_n) \in \mathbf{C}^n$. In addition, since $G = 0$ on $\mathbf{R}^n$ and G is analytic, it follows that $G = 0$ on $\mathbf{C}^n$. In particular, this means that $G(iy_1, iy_2, \dots, iy_n) = 0$

for all $y = (y_1, y_2, \dots, y_n) \in \mathbf{R}^n$. Then, for $\phi \in C_0^\infty(\mathbf{R}^n)$,

$$\begin{aligned}
&\int_\Omega \phi(\Delta W_{t_1}(\omega), \dots, \Delta W_{t_n}(\omega)) g(\omega)\, dP^0(\omega) \\
&\quad = \int_\Omega (2\pi)^{-n/2} \left(\int_{\mathbf{R}^n} \hat{\phi}(y) e^{i(y_1 \Delta W_{t_1}(\omega) + \dots + y_n \Delta W_{t_n}(\omega)}\, dy \right) g(\omega)\, dP^0(\omega) \\
&\quad = (2\pi)^{-n/2} \int_{\mathbf{R}^n} \hat{\phi}(y) \left(\int_\Omega e^{i(y_1 \Delta W_{t_1}(\omega) + \dots + y_n \Delta W_{t_n}(\omega)} g(\omega)\, dP^0(\omega) \right) dy \\
&\quad = (2\pi)^{-n/2} \int_{\mathbf{R}^n} \hat{\phi}(y) G(iy) dy = 0
\end{aligned}$$

where

$$\hat{\phi}(y) = (2\pi)^{-n/2} \int_{\mathbf{R}^n} \phi(x) e^{-i(x,y)}\, dx$$

is the Fourier transform of ϕ and we have used the inverse Fourier transform Theorem

$$\phi(x) = (2\pi)^{-n/2} \int_{\mathbf{R}^n} \hat{\phi}(y) e^{i(x,y)}\, dy.$$

Now, since $G(iy) \neq 0$ it must be $g = 0$. In addition, g is orthogonal to a dense subset of $L^2(\mathcal{F}_T, P^0)$ (Lemma 36). Therefore, the linear span of the functions (1.63) must be dense in $L^2(\mathcal{F}_T, P^0)$ as claimed.

Assume now that $f(t,\omega) \in \mathcal{L}_T^2$. Then, the random variable

$$F(\omega) \equiv \int_0^T f(t,\omega)\, dW_t(\omega)$$

is $\mathcal{F}_T$-measurable and by the Itô isometry

$$E[F^2] = E\left[\int_0^T f(t,\omega)\, dW_t(\omega)\right]^2 = \int_0^T E[f^2(t,\omega)]\, dt < \infty$$

so that $F \in L^2(\mathcal{F}_T, P^0)$.

Theorem 38. (Itô Representation Theorem) *Let $F \in L^2(\mathcal{F}_T, P^0)$. Then there exists a unique stochastic process $f(t,\omega) \in \mathcal{L}_T^2$ such that*

$$F(\omega) = E[F] + \int_0^T f(t,\omega)\, dW_t(\omega). \tag{1.66}$$

Proof. Let's start assuming that F is of the form (1.63), i.e.

$$F_t(\omega) = \exp\left\{ \int_0^t h(s)\, dW_s(\omega) - \frac{1}{2} \int_0^t h^2(s)\, ds \right\}$$

for $t \in [0,T]$ and some $h(s) \in L^2[0,T]$. Then by Itô formula and since $F_0(\omega) = 1$,

$$dF_t(\omega) = F_t(\omega) \cdot h(t)\, dW_t(\omega)$$

so that

$$F(\omega) \equiv F_T(\omega) = 1 + \int_0^T F_s(\omega) h(s)\, dW_s(\omega)$$

from which it follows that $E[F] = 1$. This shows that (1.66) holds in this case and, by linearity, also holds for linear combinations of functions of the form (1.63). If $F \in L^2(\mathcal{F}_T, P^0)$ is arbitrary, we approximate it in $L^2(\mathcal{F}_T, P^0)$ by linear combinations F_n of functions of the form (1.63) which, from Lemma 37, we know is dense in $L^2(\mathcal{F}_T, P^0)$. and therefore $F_n \to F$ pointwise a.e. P^0 as $n \to \infty$. For each n we have

$$F_n(\omega) = E[F_n] + \int_0^T f_n(s,\omega)\, dW_s(\omega)$$

and $f_n \in \mathcal{L}^2_T$. Now, using Theorem 6.b and c,

$$\begin{aligned}
E[(F_n - F_m)^2] &= E\left[E[F_n - F_m] + \int_0^T (f_n - f_m)\, dW)^2\right] \\
&= (E[F_n - F_m])^2 + 2 \cdot E\left[[F_n - F_m] \cdot \int_0^T (f_n - f_m)\, dW\right] \\
&+ E\left[\int_0^T (f_n - f_m)\, dW\right]^2 \\
&= (E[F_n - F_m])^2 + \int_0^T E[(f_n - f_m)^2]\, dt \to 0
\end{aligned}$$

as $n, m \to \infty$. This proves that $\{f_n : n \in \mathbf{N}\}$ is a Cauchy sequence in $L^2([0,T] \times \Omega)$ and hence converges to some $f \in L^2([0,T] \times \Omega)$. Since $f_n \in \mathcal{L}^2_T$ we have $f \in \mathcal{L}^2_T$ and, again using Itô isometry,

$$F = \lim_n F_n = \lim_n \left(E[F_n] + \int_0^T f_n\, dW \right) = E[F] + \int_0^T f\, dW,$$

the limit being taken in $L^2(\mathcal{F}_T, P^0)$. Hence the representation (1.66) holds for all $F \in L^2(\mathcal{F}_T, P^0)$.
The uniqueness is a consequence of Itô isometry. In fact, if we assumed that

$$F(\omega) = E[F] + \int_0^T f_1(t,\omega)\, dW_t(\omega) = E[F] + \int_0^T f_2(t,\omega)\, dW_t(\omega)$$

with $f_1, f_2 \in \mathcal{L}^2_T$, then

$$0 = E\left[\left(\int_0^T (f_1(t,\omega) - f_2(t,\omega))\, dW_t(\omega)\right)^2\right] = \int_0^T E[(f_1(t,\omega) - f_2(t,\omega))^2]\, dt$$

and therefore $f_1(t,\omega) = f_2(t,\omega)$ for a.a. $(t,\omega) \in [0,T] \times \Omega$.

Problem Q. Find the process $f(t,\omega) \in \mathcal{L}_T^2$ such that (1.66) holds if $F(\omega) = W_T^2(\omega)$.

Solution. If $F(\omega) = W_T^2(\omega)$ then $E[F] = T$. Thus, $f(t,\omega) = 2W_t(\omega)$ is the right choice. In fact, we know that $\int_0^T W_t(\omega)\, dW_t(\omega) = \frac{1}{2}W_T^2(\omega) - \frac{T}{2}$. So,

$$W_T^2(\omega) = T + \int_0^T 2W_t(\omega)\, dW_t(\omega).$$

Problem R. Find the process $f(t,\omega) \in \mathcal{L}_T^2$ such that (1.66) holds if $F(\omega) = W_T^3(\omega)$.

Solution. $f(t,\omega) = 3(W_t^2(\omega) - T + t)$. In fact, with this choice for $f(\omega,t)$ and the fact that $E[F] = 0$ we have

$$\begin{aligned}
\int_0^T 3(W_t^2(\omega) - T + t)\, dW_t(\omega) &= \int_0^T 3W_t^2(\omega)\, dW_t(\omega) - 3TW_T(\omega) + 3\int_0^T t\, dW_t(\omega) \\
&= W_T^3(\omega) - 3\int_0^T W_t(\omega)\, dt - 3TW_T(\omega) + 3\int_0^T t\, dW_t(\omega) \\
&= W_T^3(\omega) + 3TW_T(\omega) - 3\int_0^T t\, dW_t(\omega) - 3TW_T(\omega) + 3\int_0^T t\, dW_t(\omega) \\
&= W_T^3(\omega)
\end{aligned}$$

Problem S. Find the process $f(t,\omega) \in \mathcal{L}_T^2$ such that (1.66) holds if $F(\omega) = e^{W_T(\omega)}$.

Solution. $f(t,\omega) = e^{W_t(\omega)+(T-t)/2}$. First one should note that $E[F] = e^{T/2}$ and then it is easily checked that

$$dY_t(\omega) = Y_t(\omega) dW_t(\omega)$$

has the solution $Y_t(\omega) = e^{W_t(\omega)-t/2}$. This means that

$$de^{W_t(\omega)-t/2} = e^{W_t(\omega)-t/2} dW_t(\omega)$$

or,

$$e^{W_T(\omega)-T/2} = 1 + \int_0^T e^{W_t(\omega)-t/2}\, dW_t(\omega)$$

and, multiplying both sides by $e^{T/2}$,

$$e^{W_T(\omega)} = e^{T/2} + \int_0^T e^{W_t(\omega)+(T-t)/2}\, dW_t(\omega)$$

as it was claimed.

Problem T. Find the process $f(t,\omega) \in \mathcal{L}_T^2$ such that (1.66) holds if $F(\omega) = \sin W_T(\omega)$.

Solution. $f(t,\omega) = e^{(1/2)(t-T)} \cos W_t(\omega)$. In fact, using Itô formula, we have

$$d(e^{(1/2)t} \sin W_t(\omega)$$
$$= \left(\frac{1}{2}e^{(1/2)t} \sin W_t(\omega) - \frac{1}{2}e^{(1/2)t} \sin W_t(\omega)\right) dt + e^{(1/2)t} \cos W_t(\omega)\, dW_t(\omega).$$

This is the same as

$$e^{(1/2)T} \sin W_T(\omega) = 0 + \int_0^T e^{(1/2)t} \cos W_t(\omega)\, dW_t(\omega)$$

from which, given that $E[F] = 0$, our conclusion above follows easily.

Martingale Representation Theorem

Theorem 39. (The Martingale Representation Theorem) *Suppose $\{M_t, \mathcal{F}_t : t \in \mathbf{R}_+\}$ is a martingale with respect to P^0 and that $M_t \in L^2(P^0)$ for all $t \geq 0$. Then there exists a unique stochastic process $f(s,\omega)$ such that $f(t,\cdot) \in \mathcal{L}^2[0,t]$ for all $t \geq 0$ and*

$$M_t(\omega) = E[M_0] + \int_0^t f(s,\omega)\, dW_s(\omega); \quad t \geq 0.$$

Proof. From Theorem 38 (let $F = M_t, t = T$) for all t there exists a unique $f^t(s,\omega) \in L^2(\mathcal{F}_t, P^0)$ such that

$$M_t(\omega) = E[M_t] + \int_0^t f^t(s,\omega)\, dW_s(\omega) = E[M_0] + \int_0^t f^t(s,\omega)\, dW_s(\omega).$$

Now, if $0 < t_1 < t_2$, we find

$$M_{t_1} = E[M_{t_2} \mid \mathcal{F}_{t_1}] = E[M_0] + E\left[\int_0^{t_2} f^{t_2}(s,\omega)\, dW_s(\omega) \mid \mathcal{F}_{t_1}\right]$$
$$= E[M_0] + \int_0^{t_1} f^{t_2}(s,\omega)\, dW_s(\omega).$$

Clearly, we have also

$$M_{t_1} = E[M_0] + \int_0^{t_1} f^{t_1}(s,\omega)\, dW_s(\omega)$$

and comparing the last two equalities, it follows that

$$0 = E\left[\left(\int_0^{t_1} (f^{t_1}(s,\omega) - f^{t_2}(s,\omega))\, dW_s\right)^2\right] = \int_0^{t_1} E[(f^{t_1}(s,\omega) - f^{t_2}(s,\omega))^2]\, ds$$

and hence

$$f^{t_1}(s,\omega) = f^{t_2}(s,\omega) \text{ for a.a. } (s,\omega) \in [0,t_1] \times \Omega.$$

Then, it is possible to define $f(s,\omega)$ for a.a. $(s,\omega) \in [0,\infty) \times \Omega$ by setting

$$f(s,\omega) = f^N(s,\omega)$$

if $s \in [0, N]$ and this gives

$$M_t = E[M_0] + \int_0^t f^t(s,\omega)\,dW_s(\omega) = E[M_0] + \int_0^t f(s,\omega)\,dW_s(\omega)$$

for all $t \geq 0$.

Problem U. (Øksendal (1995) Problem 4.12) Let $\{\mathcal{F}_t : t \in \mathbf{R}_+\}$ be the natural filtration for a Brownian motion in $\mathbf{R}$ such that $W_0(\omega) = 0$ and

$$dX_t(\omega) = u(t,\omega)\,dt + v(t,\omega)\,dW_t(\omega)$$

an Itô process in $\mathbf{R}$ such that

$$E\left[\int_0^t |u(r,\omega)|\,dt\right] + E\left[\int_0^t v^2(r,\omega)\,dr\right] < \infty$$

for all $t \geq 0$. Suppose X_t is an $\{\mathcal{F}\}_t$-martingale. Then

$$u(s,\omega) = 0 \;\; \textit{for a.a.}\; (s,\omega) \in \mathbf{R}_+ \times \Omega.$$

Solution. Using Theorem 27 we have that for all $s \geq t$

$$\begin{aligned} E[X_s(\omega) \mid \mathcal{F}_t] &= E\left[\int_0^s u(r,\omega)\,dr \mid \mathcal{F}_t\right] + E\left[\int_0^s v(r,\omega)\,dW_r(\omega) \mid \mathcal{F}_t\right] \\ &= \int_0^t u(r,\omega)\,dr + \int_0^t v(r,\omega)\,dW_r(\omega) + E\left[\int_t^s u(r,\omega)\,dr \mid \mathcal{F}_t\right] + 0 \\ &= X_t(\omega) + E\left[\int_t^s u(r,\omega)\,dr \mid \mathcal{F}_t\right]. \end{aligned}$$

Clearly, if $\{X_t, \mathcal{F}_t : t \in \mathbf{R}_+\}$ is to be a martingale, we must have

$$E\left[\int_t^s u(r,\omega)\,dr \mid \mathcal{F}_t\right] = 0 \;\text{ for all } s \geq t.$$

If we differentiate the last equality with respect to s we find

$$E[u(s,\omega) \mid \mathcal{F}_t] = 0 \;\text{ a.s. for a.a. } s \geq t.$$

Now, using our assumptions, it follows that $E[u(s,\omega) \mid \mathcal{F}_t]$ is a u.i. martingale. This means that as an application of the Martingale Convergence Theorem we have when we let $t \uparrow s$

$$0 = E[u(s,\omega) \mid \mathcal{F}_t] \to E[u(s,\omega) \mid \mathcal{F}_s] = u(s,\omega).$$

Thus, we have established that $u(s,\omega)=0$ for a.a. $(s,\omega)\in \mathbf{R}_+\times\Omega$.

Remark. The last result holds only when we are using the filtration generated by Brownian motion, $\{\mathcal{F}_t : t\in\mathbf{R}_+\}$. It does not hold if we replace $\{\mathcal{F}_t : t\in\mathbf{R}_+\}$ by $\{\mathcal{M}_t : t\in\mathbf{R}\}$ where $\mathcal{M}_t=\sigma(X_r(\omega) : r\le t)$ and we assume that $\{X_t,\mathcal{M}_t : t\in\mathbf{R}_+\}$ is a martingale. Nevertheless, the Brownian martingale representation, also know as predictable representation, holds for local martingales and for local martingale adapted to Poisson filtration.

Problem V. Let X_t be an Itô integral

$$dX_t(\omega)=f(t,\omega)\,dW_t(\omega)$$

where $f\in\mathcal{L}^2_T$. Give an example to show that X_t^2 is not in general a martingale. On the other hand, if $f(x)$ is a bounded and continuous function on $\mathbf{R}$, Itô integrals like $\int_0^t f(W_s(\omega))\,dW_s(\omega)$ are square integrable martingales on any finite interval $[0,T]$.

Solution. Let $t>1/4$ and $f(t,\omega)=e^{W_t^2(\omega)}$. It is easily seen that $X_t=X_0+\int_0^t e^{W_s^2(\omega)}\,dW_s(\omega)$ for every $t\in[0,T]$ and

$$X_t^2=X_0^2+2X_0\cdot\int_0^t e^{W_s^2(\omega)}\,dW_s(\omega)+\left(\int_0^t e^{W_s^2(\omega)}\,dW_s(\omega)\right)^2.$$

Thus,

$$E[X_t^2]=E\left[\left(\int_0^t e^{W_s^2(\omega)}\,dW_s(\omega)\right)^2\right]=\int_0^t E[e^{2\cdot W_s^2(\omega)}]\,ds=\infty$$

as for $s>1/4$, $E[e^{2W_s^2(\omega)}]=\infty$.
The second part is easy since $|f(W_s(\omega))|\le K$ for some real K for $s\in[0,T]$. Hence,

$$E\left[\left(\int_0^t f(W_s(\omega))dW_s(\omega)\right)^2\right]\le K^2t\le K^2T<\infty.$$

The last result can be extended to the case where $\int_0^\infty f(s)\,ds<\infty$. In fact, under this assumption, $\int_0^t f(W_s(\omega))\,dW_s(\omega)$ is a square integrable martingale on $[0,\infty)$.

If $dX_t(\omega)=u(t,\omega)dt+dW_t(\omega)$ is an Itô process where $u(\cdot,\cdot)$ is bounded, we know from a previous exercise that unless $u(\cdot,\cdot)\equiv 0$ the process X_t is not an $\mathcal{F}_t$-martingale. However, it is possible to turn it into an $\mathcal{F}_t$-martingale by multiplying X_t by a suitable exponential martingale. More precisely, if we define

$$Y_t(\omega)=X_t(\omega)\cdot M_t(\omega)$$

where

$$M_t(\omega)=\exp\{-X_t-(1/2)\int_0^t u^2(r,\omega)\,dr\}$$

it is possible to show that Y_t is an $\mathcal{F}_t$-martingale. To prove this, the reader is invited to first work on the following

Problem W. Show that if

$$Z_t(\omega) = \exp\left\{-\int_0^t g(s,\omega)\,dW_s(\omega) - \frac{1}{2}\int_0^t g^2(s,\omega)\,ds\right\}$$

then $dZ_t(\omega) = -Z_t(\omega)\cdot g(t,\omega)dW_t(\omega)$.

Solution. Let $A_t(\omega) = \int_0^t g(s,\omega)\,dW_s(\omega)$ (i.e., $dA_t(\omega) = g(t,\omega)dW_t(\omega)$) and let

$$Z_t(\omega) = \exp\left\{-A_t(\omega) - \frac{1}{2}\int_0^t g^2(s,\omega)\,ds\right\}.$$

Using Itô formula one finds that

$$\begin{aligned} dZ_t(\omega) &= \left[-\frac{1}{2}g^2(t,\omega) + \frac{1}{2}g^2(t,\omega)\right] Z_t(\omega)\,dt - g(t,\omega)\cdot Z_t(\omega)\,dW_t(\omega) \\ &= -g(t,\omega)\cdot Z_t(\omega)\,dW_t(\omega) \end{aligned}$$

as we were supposed to show.

Now, using the multidimensional Itô formula, we can show that

$$\begin{aligned} dY_t(\omega) &= M_t(\omega)\cdot dX_t(\omega) + X_t(\omega)\cdot dM_t(\omega) + dX_t(\omega)\cdot dM_t(\omega) \\ &= M_t(\omega)\cdot[g(t,\omega)dt + dW_t(\omega)] - X_t(\omega)\cdot M_t(\omega)\cdot g(t,\omega)dW_t(\omega) \\ &+ [g(t,\omega)dt + dW_t(\omega)][-M_t(\omega)g(t,\omega)dW_t(\omega)] \\ &= M_t(\omega)g(t,\omega)dt + M_t(\omega)dW_t(\omega) - X_t(\omega)M_t(\omega)g_t(\omega)dW_t(\omega) \\ &- M_t(\omega)g^2(t,\omega)(dt\cdot dW_t(\omega)) - M_t(\omega)g(t,\omega)(dW_t(\omega))^2 \\ &= M_t(\omega)g(t,\omega)dt + M_t(\omega)dW_t(\omega) - X_t(\omega)M_t(\omega)g_t(\omega)dW_t(\omega) \\ &+ 0 - M_t(\omega)g(t,\omega)dt = M_t(\omega)[1 - X_t(\omega)g(t,\omega)]dW_t(\omega). \end{aligned}$$

It is easily seen that under our assumptions

$$M_s(\omega)\left[1 - \int_0^t g(s,\omega)ds + W_t(\omega)\right]$$

is in $\mathcal{L}_T^2$ and since, in addition, we have also $Y_0 = E[Y_0] = 0$ the desired conclusion follows from Theorem 39. This result is a special case of the **Girsanov Theorem** and can be interpreted as follows: $\{X_t : t \in \mathbf{R}_+\}$ is a martingale with respect to the measure Q defined on $\mathcal{F}_T$ by $dQ = M_T dP^0$ for $T < \infty$. An equivalent way of expressing this is to say that $Q << P^0$ (Q is absolutely continuous with respect to P^0) with Radon-Nikodym derivative

$$\frac{dQ}{dP^0} = M_T \text{ on } \mathcal{F}_T.$$

In particular, since $M_T(\omega) > 0$ a.s., we have also $P^0 << Q$ and hence P^0 and Q are equivalent measures.

Theorem 40. (Girsanov Theorem I) *Consider a probability measure P on the space of paths $\{B_t(\omega),\ t \le T\}$ such that $B_t(\omega)$ is a Brownian motion and assume that $b_t(\omega)$ is a predictable function for every $t \in [0,T]$. Set* Girsanov Theorem I

$$M_t(\omega) = \exp\left\{ \int_0^t b(s,\omega)\, dB_t(\omega) - (1/2) \int_0^t (b(s,\omega))^2\, ds \right\} \quad t \le T,$$

where $b(t,\omega)$ satisfies the Novikov's condition

$$E\left[\exp\left\{ \frac{1}{2} \int_0^T b(s,\omega)^2\, ds \right\} \right] < \infty$$

and define a new measure Q on the set of trajectories $\{B_(t\omega) : t \le T\}$ by

$$Q(\{\delta\}) \equiv E^P[\{\delta\}],$$

where δ represents an arbitrary set of paths and E^P is the expectation operator with respect to the probability P. Then, the random process

$$W_t(\omega) = B_t(\omega) - \int_0^t b(s,\omega)\, ds \quad t \le T,$$

is a Brownian motion under the measure Q.

Proof. The basic idea of the proof is to show that

$$E_t^Q[\{e^{\lambda(W_T(\omega) - W_t(\omega))}\}] = e^{\frac{\lambda^2}{2}(T-t)}. \tag{1.67}$$

In fact, if this were the case, then we have established that $W_T - W_t$ is Gaussian with mean zero and variance $T - t$. In addition, (1.67) also implies that two successive increments, say $W_{t+a} - W_t$ $W_{t+a+b} - W_{t+a}$ are statistically independent. Since gaussianity and independence of increments characterize Brownian motion, establishing (1.67) is all one has to do. Consider first the case $t = 0$. Then,

$$\begin{aligned}
E^Q[e^{\lambda W_T(\omega)}] &= E^P[e^{\lambda W_T(\omega)} M_T(\omega)] \\
&= E^P\left[e^{\lambda\left(B_T(\omega) - \int_0^T b(s,\omega)\, ds\right) + \int_0^T b(s)\, dB_s - (1/2)\int_0^T (b(s,\omega))^2\, ds} \right] \\
&= E^P\left[\exp\left\{ \int_0^T (\lambda + b(s,\omega))\, dB_s(\omega) - \int_0^T \lambda b(s,\omega) - \frac{1}{2} \int_0^T (b(s,\omega))^2\, ds \right\} \right] \\
&= E^P\left[\exp\left\{ \int_0^T (\lambda + b(s,\omega))\, dB_s(\omega) - \frac{1}{2} \int_0^T (\lambda + b(s,\omega))^2\, ds + \frac{1}{2}\lambda^2 T \right\} \right] \\
&= E^P\left[\exp\left\{ \int_0^T (\lambda + b(s,\omega))\, dB_s(\omega) - \frac{1}{2} \int_0^T (\lambda + b(s,\omega))^2\, ds \right\} \right] \cdot \exp\left\{ \frac{1}{2}\lambda^2 T \right\}
\end{aligned}$$

and, since $H_t \equiv \exp\left\{\int_0^T(\lambda + b(s,\omega))\,dB_s(\omega) - \frac{1}{2}\int_0^T(\lambda + b(s,\omega))^2\,ds\right\}$ is a martingale (under Novikov's condition) such that $H_0 = 1$ a.s., it follows that its expectation is also 1. This proves that

$$E^Q[e^{\lambda W_T}] = e^{\frac{\lambda^2}{2}T}$$

as we were supposed to show. We have established (1.67) for $t = 0$, but the calculation of the conditional moment-generating function of the increments $W_{t+a} - W_t$ is analogous but we need to use the following

$$Q_t(\{\delta\}) \equiv E_t^P\left[\delta \cdot \frac{M_T}{M_t}\right].$$

To see why it is so, one should start by noticing that if we let $E_t[\cdot]$ be the conditional expectation operator with respect to the history of paths up to time t and X is an arbitrary random variable, then

$$E[XY] = E[E_t[X]Y]$$

for all random variables Y which are measurable with respect to the past up to time t. Therefore,

$$\begin{aligned} E^Q[XY] &= E^P[XYM_T] \\ &= E^P\left[X \cdot \frac{M_T}{M_t} \cdot M_t \cdot Y\right] \\ &= E^P\left[E_t^P\left[X \cdot \frac{M_T}{M_t}\right] M_t Y\right] \\ &= E^Q\left[E_t^P\left[X \cdot \frac{M_T}{M_t}\right] Y\right] \end{aligned}$$

A simple comparison between

$$E[XY] = E[E_t[X]Y] \text{ and } E^Q[XY] = E^Q[E_t^P[X\frac{M_T}{M_t}Y]$$

gives that

$$E_t^Q[X] = E_t^P\left[X \cdot \frac{M_T}{M_t}\right]$$

which proves our claim. So, the conditional probability Q_t is given explicitly by

$$Q_t(\{\delta\}) \equiv E_t^P\left[\delta \cdot \frac{M_T}{M_t}\right] = E_t^P\left[\delta \cdot \exp\left\{\int_0^T b(s,\omega)\,dB_s - \frac{1}{2}\int_t^T (b(s,\omega))^2\,ds\right\}\right]$$

where δ represents again a set of paths.

Remark. In many books the Theorem above is given in a different form. Namely, it is assumed that

$$M_t = \exp\left\{-\int_0^t b(s,\omega)\,dB_s(\omega) - \frac{1}{2}\int_0^t b(s,\omega)^2\,ds\right\}; \quad t \le T.$$

In that case we need to change the definition of W_t as follows:

$$W_t(\omega) = B_t(\omega) + \int_0^t b(s,\omega)\,ds.$$

Girsanov Theorem II

Theorem 41. **(Girsanov Theorem II)** *Let $Y_t(\omega) \in \mathbf{R}^n$ be an Itô process of the form*

$$dY_t(\omega) = \beta(t,\omega)\,dt + \theta(t,\omega)\,dB_t(\omega); \quad t \le T$$

where $B_t(\omega) \in \mathbf{R}^m$, $\beta(t,\omega) \in \mathbf{R}^{n\times m}$. Suppose there exist $\mathcal{L}_T^2$-processes $u(t,\omega) \in \mathbf{R}^m$ and $\alpha(t,\omega) \in \mathbf{R}^n$ such that

$$\theta(t,\omega)\cdot u(t,\omega) = \beta(t,\omega) - \alpha(t,\omega)$$

and assume that $u(t,\omega)$ satisfies the Novikov's condition

$$E\left[\exp\left\{\frac{1}{2}\int_0^T u(s,\omega)^2\,ds\right\}\right] < \infty.$$

Put

$$M_t(\omega) = \exp\left\{-\int_0^t u(s,\omega)\,dB_s(\omega) - \frac{1}{2}\int_0^t u^2(s,\omega)\,ds\right\}; \quad t \le T$$

and

$$dQ(\omega) = M_T(\omega)dP^0(\omega) \ \text{ on } \mathcal{F}_T.$$

Then

$$\hat{B}_t(\omega) \equiv B_t(\omega) + \int_0^t u(s,\omega)\,ds; \quad t \le T$$

is a Brownian motion with respect to Q and in terms of $\hat{B}_t$ the process Y_t has the stochastic integral representation

$$dY_t(\omega) = \alpha(t,\omega)\,dt + \theta(t,\omega), d\hat{B}_t(\omega). \tag{1.68}$$

Proof. That $\hat{B}_t$ is a Brownian motion with respect to Q follows from Theorem 40. The representation (1.68) can be proved using the argument below

$$\begin{aligned} dY_t(\omega) &= \beta(t,\omega)\,dt + \theta(t,\omega)\,dB_t(\omega) \\ &= \beta(t,\omega)\,dt + \theta(t,\omega)[d\hat{B}_t(\omega) - u(t,\omega)\,dt] \\ &= [\beta(t,\omega) - \theta(t,\omega)u(t,\omega)]\,dt + \theta(t,\omega)\,d\hat{B}_t(\omega) \\ &= \alpha(t,\omega)\,dt + \theta(t,\omega)\,d\hat{B}_t(\omega). \end{aligned}$$

12. Distributions of Stochastic Integrals and Diffusion Processes

The task of computing the distribution of stochastic integrals is in general a difficult one. Nonetheless, for some simple cases, the computation of closed form distributions is not too hard. The next example is very important as it is representative of the only family of stochastic integrals whose distribution functions are easily available.

Example. Let $\{W_t, t \in \mathbf{R}_+\}$, be a Wiener Process with parameter σ. Let α and β be finite numbers and let g be a continuously differentiable function on $[\alpha, \beta]$. Then,

$$\int_\alpha^\beta g(t)\, dW_t \sim N\left[0, \sigma^2 \cdot \int_\alpha^\beta g^2(t)\, dt\right].$$

In fact, when f is deterministic it is easily checked that

$$\int_\alpha^\beta g(t)\, dW_t = g(\beta)W_\beta - g(\alpha)W_\alpha - \int_\alpha^\beta g'(t) W_t\, dt. \tag{1.69}$$

Computing the distribution of the stochastic integral is therefore equivalent to computing the distribution of (1.69). Thus, normality follows easily since by the properties of Brownian motion and the fact that, as proved in section 2 of this Chapter when discussing random integrals, $\int_\alpha^\beta g'(t)W(t)dt$ has also a normal distribution, our stochastic integral is a sum of normal random variables. The mean is zero as

$$E[\int_\alpha^\beta g(t)dW_t] = g(\beta)E[W_\beta] - g(\alpha)E[W_\alpha] - \int_\alpha^\beta g'(t)E[W_t]dt = 0$$

since the existence of the integral makes it is possible (by Fubini's Theorem) to interchange the operators E and $\int$.
To compute the variance, we use Itô isometry: i.e.:

$$E[\int_\alpha^\beta g(t)dW_t]^2 = \sigma^2 \cdot \int_\alpha^\beta g^2(t)\, dt.$$

The example we just worked out shows that whenever the integrand is not a random function, as long as g is a "nice" function, it is not too difficult to compute the distribution of stochastic integrals. However, when g is a stochastic process the situation is, in principle, hopeless other than in just a few cases.

Example. Let $\{W_t, t \in \mathbf{R}_+\}$, be a Wiener Process with parameter $\sigma = 1$. Then,

$$\int_0^1 W_t\, dW_t \sim \frac{1}{2}(\chi^2(1) - 1).$$

This follows very easily from being (see Problem E, with $T = 1$)

$$\int_0^1 W_t\, dW_t = \frac{1}{2}[W_1^2 - 1]$$

and the properties of Brownian motion.

Moments of stochastic integrals are important too. In particular, they are important when one has to compute mean and variance of predictions in the case of diffusion processes as we will shortly see. Again, the task can be relatively easy or very hard depending on the nature of the integrand function.

Example. Let W_t, $0 \leq t < \infty$, be the Wiener Process with parameter σ. Let $X = \int_0^1 t\, dW_t$ and $Y = \int_0^1 t^2\, dW_t$. Then, $E[X] = 0$ and $E[Y] = 0$ while the correlation between X and Y equals $\sqrt{15}/4$. To check this it suffices to use the integration by parts formula; i.e.:

$$E[X] = E\left[\int_0^1 t\, dW_t\right] = E\left[tW_t \mid_0^1 - \int_0^t W_t\, dt\right]$$
$$= E[1 \cdot W_1 - \int_0^1 W_t\, dt] = 0.$$

and, using the result just established,

$$E[Y] = E\left[\int_0^1 t^2\, dW_t\right] = E\left[t^2 W_t \mid_0^1 -2 \int_0^1 t\, W_t\, dt\right]$$
$$= EE[1 \cdot W_1 - \int_0^1 tW_t\, dt] = 0 + E\left[\int_0^1 t\, dW_t\right] = 0.$$

Using Itô isometry it is easy to verify the following:

$$E[X^2] = \sigma^2 \int_0^1 t^2\, dt = \frac{\sigma^2}{3} \text{ and } E[Y^2] = E[Y^2] = \sigma^2 \int_0^1 t^4\, dt = \frac{\sigma^2}{5}.$$

Finally, using the simple fact according to which, for sufficiently regular (e.g. continuous) functions, f and g,

$$E[\int_0^1 g(t)\, dW_t \cdot \int_0^1 h(t)\, dW_t] = \sigma^2 \cdot \int_0^1 g(t) \cdot h(t)\, dt$$

it follows easily that

$$\text{Cov}[X, Y] = \sigma^2 \cdot \int_0^1 t^3\, dt = \frac{\sigma^2}{4}.$$

Finally,

$$\text{Corr}[X, Y] = \frac{\sigma^2/4}{(\sigma^2/2 \cdot \sigma^2/3)^{1/2}} = \frac{\sqrt{15}}{4}.$$

The distributions of stochastic integrals are useful in determining properties related to the solution of specific stochastic differential equations. These solutions are also known as **diffusion processes**, a concept we will introduce and work with in a different set of notes, and are a special class of stochastic processes. Not too surprisingly, Brownian motion is the basic ingredient in deriving several other stochastic processes which have proved useful in applications.

Example. The **arithmetic Brownian motion** with drift α and volatility σ which can be defined as a stochastic differential, i.e.: arithmetic Brownian motion

$$dX_t = \alpha\, dt + \sigma\, dW_t.$$

This process has the following properties:

(1) X may be positive or negative.

(2) If $u > t$, then X_u is a future value of the process relative to time t. The distribution of X_u given X_t is normal with mean $X_t + \alpha(u-t)$ and standard deviation $\sigma\sqrt{u-t}$.

(3) The variance of the forecast X_u tends to infinity as u does (given t, X_t).

To show that the second statement about X_u given X_t is correct, it suffices to write X_u as

$$X_u = X_t + \alpha \int_t^u dt + \sigma \int_t^u dW_t$$

and use the fact that $\int_t^u dW_t \sim N(0, u-t)$.

Example. The **geometric Brownian motion** with drift α and volatility σ is a stochastic process whose stochastic differential is given by geometric Brownian motion

$$dX_t = \alpha X_t\, dt + \sigma X_t\, dW_t.$$

It has the following properties:

(1) If X starts at a positive value, it will remain positive. In particular, $X_t \to \infty$ a.s. as $t \to \infty$ if $\lambda > \sigma^2/2$, $X_t \to 0$ a.s. as $t \to \infty$ if $\lambda < \sigma^2/2$ and X_t keeps oscillating between 0 and ∞ as $t \to \infty$ if $\lambda = \sigma^2/2$.

(2) X has an absorbing barrier at 0.

(3) $X_t \mid X_s$ is lognormally distributed with mean $X_s \cdot e^{\alpha(t-s)}$ and variance $X_s^2 e^{2\alpha(t-s)} \cdot \left(e^{\sigma^2(t-s)} - 1\right)$. This shows that both the mean of the forecast tend to zero when $\alpha < 0$ and to infinity when $\alpha > 0$. Similarly, the variance of the forecast tends to infinity as t does when $\alpha \geq \sigma^2/2$ and to zero when $\alpha < \sigma^2/2$.

(4) The variance of the forecast X_u tends to infinity as u does.

Property (1) is not evident. But, if one assumes that X_0 is positive and solves the stochastic differential equation, the solution turns out to be:

$$X_t = X_0 \cdot \exp\{\alpha t + \sigma W_t\}$$

which is clearly always positive. Then, letting $Y_t = \log(X_t)$ it is easily seen, using Itô formula, that when $\sigma^2/2 \neq \lambda$

$$\lim_{t\to\infty} \frac{Y_t}{t} = \left(\alpha - \frac{\sigma^2}{2}\right) + \sigma \lim_{t\to\infty} \frac{W_t}{t} = \left(\alpha - \frac{\sigma^2}{2}\right)$$

as by the law of the iterated logarithm we know that $\lim_{t\to\infty} \sup \frac{W_t}{\sqrt{2t\log\log t}} = 1$ and $\lim_{t\to\infty} \inf \frac{W_t}{\sqrt{2t\log\log t}} = -1$. If $\sigma^2/2 = \lambda$, then it is

$$\lim_{t\to\infty} \sup \frac{Y_t}{t} = \sigma \text{ and } \lim_{t\to\infty} \inf \frac{Y_t}{t} = -\sigma$$

from which the desired conclusions follow.
Property (2) does not require any comment. As for property (3), if one lets $Y_t = \ln(X_t)$, a simple application of Itô's formula gives that $dY_t = (\alpha - (1/2)\sigma^2)\,dt + \sigma\,dW_t$ from which it is easily seen that

$$Y_t \mid Y_s \sim N\Big(Y_s + \big(\alpha - \frac{\sigma^2}{2}\big)(t-s), \sigma^2(t-s)\Big).$$

Finally, since $X_t = \exp\{Y_t\}$, we use the fact for which if a random variable $Z \sim N(\mu, \sigma^2)$ then $V = \exp\{Z\} \sim \text{Lognormal}(\mu, \sigma^2)$ and $E[V] = \exp\{\mu + \sigma^2/2\}$, $\text{Var}[V] = \exp\{2(\mu + \sigma^2)\} - \exp\{2\mu + \sigma^2\}$.

mean-reverting Ornstein Uhlenbeck

Example. The **mean-reverting Ornstein-Uhlenbeck process** can be written as a stochastic differential of the form

$$dX_t = \alpha(X_t - m)\,dt + \sigma\,dW_t, \quad \alpha < 0.$$

(1) X maybe positive or negative.

(2) The conditional distribution of X_u given X_0 is normal with mean given by $m + (X_0 - m)e^{\alpha u}$ and standard error $\frac{\sigma}{\sqrt{-2\alpha}} \cdot \sqrt{(1 - e^{2\alpha u})}$.

(3) The variance of $X_u \mid X_0$ is increasing in u but tends to a finite number as u goes to infinity; the average of $X_u \mid X_0$, on the other hand is decreasing in u and tends to m as u goes to infinity.

The normality of the distribution of X_u follows easily from the normality of $\int_0^u dW_s$. To compute the mean and variance of this distribution the following technical trick can be used. Let $X_u = Q_u e^{-u} + m$. Then, Itô's formula gives $dX_u = -Q_u e^{-u}\,du + e^{-u}\,dQ_u = (m - X_u)\,dt + e^{-u}\,dQ_u$. Since the drift terms match, one can compare the volatility terms, i.e. it must be $\sigma\,dW_u = e^{-u}\,dQ_u$ which implies that $dQ_u = \sigma e^u\,dW_u$. Simple algebraic manipulations give then

$$X_u = m + (X_0 - m)e^{-u} + \sigma \int_0^u e^{s-u}\,dW_s.$$

From this expression, it is easy to compute the mean and variance of the process. In fact, since $E[\int_0^u e^{s-u}\,dW_s] = 0$ and $E[\int_0^u e^{s-u}\,dW_s]^2 = \int_0^u e^{2(s-u)}\,dt$ using properties of Itô integrals.

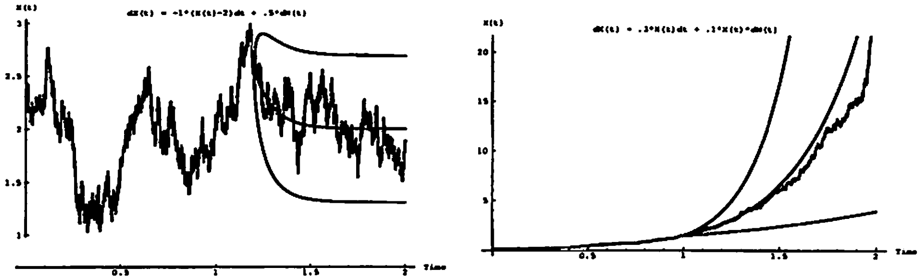

Figure 1.4: **Left.** *Mean reverting Oernstein-Uhlenbeck process with $X_0 = 2.3$, $m = 2$, $\sigma = .5$ and $\alpha = -1.0$. Superimposed to the plot are the mean function and the 95% pointwise confidence interval for $X_u \mid X_t$ when $t = 1.2$ and $u \in (1.2, 2]$.* **Right.** *Geometric Brownian motion with $X_0 = .1$, $\alpha = .3$, $\sigma = .1$. Superimposed to the plot are the mean function and the 95% pointwise confidence interval for $X_u \mid X_t$ when $t = 1$ and $u \in (1, 2]$.*

Remark. Mean reverting processes are found to be appropriate for modeling interest rates when they are known to have stable long-run values. Sometimes, volatility of stock prices is also modeled as a mean reverting process. A good introduction to modeling stochastic interest rate through diffusion models is available in Shiryaev (1999), chapter 3 section 4. A good reference for applications of such models is Rebonato (1996).

Example. This example is a little bit more difficult. We want to determine the distribution of $X_t \mid X_s$ when X_t is known to satisfy the following stochastic differential:

$$dX_t = \left(2\beta X_t + \frac{\sigma^2}{4}\right) dt + \sigma\sqrt{X_t}\, dW_t.$$

A simple inspection of the differential does not suggest an easy answer. However, there is a general result that we are going to use and, namely, if

$$dX_t = \left(\beta g(X_t)h(X_t) + \frac{\sigma^2}{2} g(X_t)g'(X_t)\right) dt + \sigma g(X_t)\, dW_t \tag{1.70}$$

where $h(x) = \int^x \frac{dy}{g(y)}$ and h has an inverse, then the solution of (1.70) is given by

$$X_t = h^{-1}\left(e^{\beta t}h(X_0) + \sigma \int_0^t e^{-\beta s}\, dW_s\right).$$

This can be proved very easily if, starting from (1.70), we let $Y_t = h(X_t)$. In fact, Itô's formula gives then

$$dY_t = \left(\beta g(x_t)h(X_t)\frac{1}{g(X_t)} + \frac{\sigma^2}{2}g(X_t)g'(X_t)\frac{1}{g(X_t)}\right.$$
$$\left. - \frac{\sigma^2}{2}g^2(X_t)\frac{g'(X_t)}{g^2(X_t)}\right)dt + \left(\sigma g(X_t)\frac{1}{g(X_t)}\right)dW_t = \beta Y_t\,dt + \sigma dW_t.$$

The last stochastic differential equation has the solution

$$Y_t = e^{\beta t}\left(Y_0 + \sigma\int_0^T e^{-\beta s}\,dW_s\right)$$

which can be verified using again Itô's formula. Finally,

$$X_t = h^{-1}(Y_t) = h^{-1}\left(e^{\beta t}h(X_0) + \sigma\int_0^t e^{-\beta s}\,dW_s\right).$$

The original stochastic differential can be written as in (1.70) if we let $g(X_t) = \sqrt{X_t}$ (and this automatically implies that $h(X_t) = 2\sqrt{X_t}$). Thus, we have that the solution of the original equation is given by

$$X_t = \left(e^{2\beta t}\sqrt{X_0} + \sigma\int_0^t e^{-2\beta s}\,dW_s\right)^2.$$

As we already know, $\int_0^t e^{-2\beta s}\,dW_s \sim N(0, \int_0^t e^{-4\beta s}\,dt$, i.e. $N(0, \frac{1}{4\beta}(1-e^{-4\beta t})$, so it is easily checked that, when X_0 is given and nonnegative,

$$e^{2\beta t}\sqrt{X_0} + \sigma\int_0^t e^{-2\beta s}\,dW_s \sim N\left(e^{2\beta t}\sqrt{X_0} + \frac{\sigma^2}{4\beta}(1-e^{-4\beta t})\right).$$

This implies that

$$X_t \mid X_0 \sim \chi^2_{(1,\delta)} \text{ where } \delta = \frac{2\beta e^{4\beta t}X_0}{\sigma^2(1-e^{-4\beta t})}.$$

Using properties of the non-central chi-squared distribution, it is also found that

$$E[X_t \mid X_0] = 1 + \frac{4\beta e^{4\beta t}X_0}{\sigma^2(1-e^{-4\beta t})} \text{ and } \mathrm{Var}[X_t \mid X_0] = 2 + \frac{16\beta e^{4\beta t}X_0}{\sigma^2(1-e^{-4\beta t})}$$

which shows that both the forecast and its standard error go to infinity as t does.

Example. The following Itô stochastic differential

$$dX_t = 1\,dt + 2\sqrt{X_t}\,dW_t$$

is not so easy to find the distribution for. But, we know that a solution X_t to the previous stochastic differential equation is also a solution to the following Fisk-Stratonovich stochastic differential equation (1.53)

$$dX_t = 0\,dt + 2\sqrt{X_t}\circ dW_t.$$

This is equivalent to $X_t^{-1/2}dX_t = 2 \circ W_t$ whose solution is easily seen to be $2\sqrt{X_t} = 2\sqrt{X_0} + 2W_t$ or $X_t = (\sqrt{X_0} + W_t)^2$ and hence $X_t \mid X_0$ is distributed as a noncentral chi-squared r.v. with noncentrality parameter $2X_0^2/t$.

Modern financial theory, derivative theory in particular, is based on the random movements of financial quantities. Deterministic equations for the values of these quantities can be derived when the building block is the Wiener process. Unfortunately, deterministic equations are seldom satisfactory in describing how the future values of financial quantities will evolve. For example, what is the probability of a call or put option to expire in the money? Deterministic equations can tell us a few things about the values of these options but not about their probability to expire in the money. Diffusion models, transition probabilities density functions, first exit times and steady state distributions are the right tools for this kind of problems.

The few examples above have shown how to find the distributions of some diffusion processes. Our approach has been that of explicitly solving the Itô stochastic differential equations (more properly, stochastic integral equations) that are used to define them. Unfortunately, this is possible for a small number of cases only. To compute the distributions for general diffusion processes one needs to use a different approach, namely, the general theory of Markov processes, and Kolmogorov backward and forward equations. Wilmott (1998), chapter 10 is a suitable first introduction to this type of problems.

Using Girsanov Theorem and its generalizations (Liptser, Shiryaev (1977), chapter 6) and the absolute continuity of measures corresponding to Itô processes and processes of diffusion type (Liptser, Shiryaev (1977), chapter 7) one can compute the characteristic functions of quadratic or bilinear functional of the Brownian motion. This fact has important implications when dealing with the statistical properties of regression models for nonstationary time series. Several important examples relative to the distribution of quadratic functionals of Brownian motion and Ornstein-Uhlenbeck processes are available in Tanaka's (1996) book.

13. References

Borkar, V. S., *Probability Theory. An Advanced Course*, Springer 1995.

Brockwell, P. J, Davis, R. A., *Introduction to Time Series and Forecasting*, Springer, 1996 (Section 10.4 contains an introduction to continuous-time models)

Fleming, T. R., Harrington, D. P., *Counting processes and survival analysis*, Wiley, 1991.

Folland, G. B., *Real Analysis, Modern Techniques and its Applications*, Wiley 1986.

Friedman, A., *Stochastic Differential Equations and Applications, Vol. I*, Academic Press 1975.

Hamilton, J. D., *Time Series Analysis*, Princeton University Press, 1994

Harrison, J. M., Pliska, S. R., *Martingales and Stochastic Integrals in the Theory of Continuous Trading*, Stochastic Processes and their Applications, **11** (1981) 215-60.

Harrison, J. M., Pliska, S. R., *A Stochastic Model of Continuous Trading*, Stochastic Processes and their Applications, **15** (1983) 313-6

Jones, F., *Lebesgue Integration on Euclidean Space*, Jones and Bartlett, 1993.

Kallianpur, G. *Stochastic Filtering Theory*, Springer 1980.

Klebaner, F. C., *Introduction to Stochastic Calculus with Applications*, Imperial College Press 1998.

Kloeden, P. E., Platen, E., *Numerical Solution of Stochastic Differential Equations*, Springer, 1992.

Liptser, R. S., Shiryaev, A. N. *Statistics of Random Processes I. General Theory*, Springer and Verlag 1977.

Mao, X., *Stochastic Differential Equations and their Applications*, Horwood Pub., 1997.

Øksendal, B., *Stochastic Differential Equations, 4th Edition*, Springer 1995.

Protter, P., *Stochastic Integration and Differential Equations*, Springer 1990.

Rebonato, R. *Interest-Rate Option Models. Understanding, Analysing and Using Models for Exotic Interest Rate Options*, Wiley 1996

Shiryaev, A. N., *Essentials of Stochastic Finance. Facts, Models, Theory*, World Scientific Publishing Co. 1999

Tanaka, K., *Time Series Analysis; Nonstationarity and Noninvertible Distribution Theory*, Wiley 1996

Todorovic, P., *An Introduction to Stochastic Processes and Their Applications*, Springer-Verlag 1992.

Wilmott, P., *Derivatives. The Theory and Practice of Financial Engineering*, Wiley 1998